The Journal of Experimental Education

A periodical report of scientific investigations relating to child development, curriculum, learning, teaching, supervision, measurements, statistics, and experimental techniques.

Vol. XXVIII | March 1960 | No. 3

CONTENTS

PAGE

Multiple Factor Analysis of Academic Achievement: A Comparative Study of Full-Blooded Indian and White Children — *Omer John Rupiper* 177

Attitude of Kindergarten Children Toward Specially Prepared Story Materials — *Bill Snider* 207

Analysis of MMPI Data for Classification Purposes by Multivariate Statistical Techniques—Peter P. Rempel 219

On Programming the Integral Rth Root of a Number on Friden Standard and Square Root Desk Calculators — *William J. Moonan* 229

Must All Tests be Multi-Factor Batteries? — *Edwin F. Rosinski* 235

An Evaluation of the Laboratory in a College General Biology Course — *Douglas M. Dearden* 241

Application of Matrices in the Analysis of Sociometric Data — *John X. Jamrich* 249

A Table to Facilitate the Calculation of Normalized T-Scores for Small Samples (N> 5 <51) — *Peter F. Merenda* 253

Estimation of Test Reliability by Analysis of Variance Technique — *Cyril J. Hoyt and P. R. Krishnaiah* 257

On Teaching F Instead of t — *Norman H. Anderson* 261

Computation of Variance Accounted for in Multiple Correlation — *Clinton I. Chase* 265

On a Test of Whether Two Sets of Observations Belong to a Certain Class of Distributions When the Random Variables are not Independent — *H. P. Kuang* 267

$7.50 A YEAR | PUBLISHED QUARTERLY | $1.90 A COPY

Published by Dembar Publications, Inc.,
Madison 3, Wisconsin.
Second class postage paid at Madison, Wisconsin

EDITORIAL BOARD

A. S. Barr, Chairman, Professor of Education, University of Wisconsin, Madison 6, Wis.

Jacob O. Bach, Director, Educational Research Services, Southern Illinois University, Carbondale, Illinois. Editorially responsible for materials on university field research.

Arthur T. Jersild, Professor of Education, Teachers College, Columbia University, New York City. Editorially responsible for materials on child welfare, guidance, and development, published each December.

Palmer O. Johnson, Professor of Education, University of Minnesota, Minneapolis, Minnesota. Editorially responsible for materials on measurements, statistics, and methods of experimental research, published each March.

H. H. Remmers, Professor of Educational Psychology, Director Division Educational Reference, Purdue University, Lafayette, Indiana. Editorially responsible for materials on learning, teaching and supervision, published each September.

J. Wayne Wrightstone, Director, Bureau of Educational Research, Board of Education of the City of New York, Brooklyn, New York, 110 Livingston Street, Brooklyn, New York. Editorially responsible for materials on curriculum construction, published each June.

CONTRIBUTING EDITORS

Emmett A. Betts, Director, Betts Reading Clinic, Haverford, Pennsylvania.

Leo J. Brueckner, Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Oscar K. Buros, Associate Professor of Education, Rutgers University, New Brunswick, New Jersey.

Guy T. Buswell, Professor of Educational Psychology, University of California, Berkeley, California.

Harold D. Carter, Associate Professor of Education, University of California, Berkeley 4, California.

Leslie L. Chisholm, Associate Professor of Education, State College of Washington, Pullman, Washington.

Herbert S. Conrad, Technical Consultant, College Entrance Examination Board, Princeton, New Jersey.

Stephen M. Corey, Professor of Education, Teachers College, Columbia University, New York, New York.

Robert A. Davis, Professor of Education, George Peabody College for Teachers, Nashville, Tennessee.

Hari R. Douglass, Director of College of Education, University of Colorado, Boulder, Colorado.

Harold A. Edgerton, Director, Occupational Opportunities Service, Professor of Psychology, Ohio State University, Columbus 10, Ohio.

John C. Flanagan, Professor of Psychology, University of Pittsburgh, Pennsylvania.

Carter V. Good, Dean, Teachers College, University of Cincinnati, Cincinnati 21, Ohio.

Robert W. B. Jackson, Assistant Professor of Educational Research; Assistant Director, Department of Educational Research, Ontario College of Education, University of Toronto, Toronto, Canada.

Harold E. Jones, Professor of Psychology and Director, Institute of Child Welfare, University of California, Berkeley 4, California.

Noel Keys, Professor of Education and Lecturer in Human Relations, University of California, Berkeley, California.

D. Welty Lefever, Professor of Education, University of Southern California, Los Angeles, California.

Edward A. Lincoln, Consulting Psychologist, Halifax, Massachusetts.

Irving Lorge, Professor of Education, Executive Officer, Institute of Psychological Research, Teachers College, Columbia University, New York 27, New York.

A. R. Mead, Director of Educational Research, University of Florida, 330 P. K. Younge Building, Gainesville, Florida.

T. E. Newland, Professor of Education, University of Illinois, Urbana, Illinois.

C. W. Odell, Professor of Education, University of Illinois, Urbana, Illinois.

Willard C. Olson, Dean, School of Education, University of Michigan, Ann Arbor, Michigan.

Valworth R. Plumb, Chairman, Division of Education and Psychology, University of Minnesota (Branch), Duluth, Minnesota.

S. L. Pressey, Professor of Educational Psychology, Ohio State University, Columbus, Ohio.

William Reitz, Associate Professor of Education, College of Education Examiner, Wayne University, Detroit 2, Michigan.

Henry D. Rinsland, Professor of Education and Director of Educational Research, The University of Oklahoma, Norman, Oklahoma.

Robert T. Rock, Jr., Professor of Psychology, Head of Dept. of Psychology, Graduate School, Fordham University, New York City.

Philip J. Rulon, Professor of Education, Harvard Graduate School of Education, Cambridge 38, Massachusetts.

John Schmid, Research and Evaluation Psychologist, Lackland Air Force Base, San Antonio, Texas.

Louis G. Schmidt, Assistant Professor of Counseling Psychology, School of Education, Indiana University, Bloomington, Indiana.

Harold Seashore, Director, Test Division, The Psychological Corporation, New York 18, New York.

David Segel, Educational Consultant, Specialist in Tests and Measurements, Federal Security Agency, U. S. Office of Education, Washington, D. C.

Paul W. Terry, Professor of Educational Psychology, University of Alabama, University, Alabama.

Helen Thompson, Associate Attending Psychologist, New York Post-Graduate Hospital, 303 East 20th Street, New York 3, N. Y.

Robert L. Thorndike, Associate Professor of Education, Teachers College, Columbia University, New York City.

Herbert A. Toops, Professor of Psychology, Ohio State University, Columbus, Ohio.

Maurice E. Troyer, Vice President, Japan International Christian University, Tokyo, Japan.

Helen M. Walker, Professor of Education, Teachers College, Columbia University, New York City.

Beth L. Wellman, Professor of Psychology, Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

Guy M. Wilson, Emeritus Professor of Education, Boston University, 33 Pine Street, Wellesley Hills, Massachusetts.

Paul A. Witty, Professor of Education, Director of Psycho-Educational Clinic, School of Education, Northwestern University, Evanston, Illinois.

Ernest R. Wood, Professor of Education, New York University, New York City.

D. A. Worcester, Chairman Department of Educational Psychology and Measurement, University of Nebraska, Lincoln, Nebraska.

DEMOCRAT PRINTING COMPANY
MADISON, WISCONSIN

The Journal of Experimental Education

PUBLISHED QUARTERLY

Copyright 1960 by

DEMBAR PUBLICATIONS, INC.
Madison 3, Wisconsin

JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 28, Number 3, March 1960)

MULTIPLE FACTOR ANALYSIS OF ACADEMIC ACHIEVEMENT: A COMPARATIVE STUDY OF FULL-BLOODED INDIAN AND WHITE CHILDREN*

OMER JOHN RUPIPER**
University of Kansas

I. Introduction

THIS STUDY had its origin in the studies conducted by the Bureau of Indian Affairs and the University of Kansas during the period 1951-55 (3, 13). This Indian Education Research Project is undoubtedly the most comprehensive study of Indian and white children which has been made. The program explored the educational status of Indian children and emphasized two functions: predictive testing and achievement testing. The purpose of the predictive testing was an attempt to investigate high-school graduates' academic skills and to predict their probable success if they consider further formal education. The achievement testing was designed to serve two purposes: (a) a continuing evaluation of the status of educational achievement of Indian children for administrative use, that is, the recording and interpreting of test scores for describing a pupil's educational achievement as measured by tests, and (b) the use of data obtained as a functional part of the instructional program. These studies showed that Indian children differed from white children as measured by achievement test results and that the smaller the proportion of full-blooded and non-English speaking children in a specific group of Indian children the higher the group achieved. It was recognized that these differences in educational achievement depended on differences in experiential, environmental, and cultural factors and not necessarily on innate ability alone.

The present study was an attempt to isolate and compare the factor structures and sets of factor patterns obtained from the achievement test scores for full-blooded Indian children and white children. The purpose of this study was to investigate the degree of relationship between the subtests of the measuring instrument, to determine the number of independent factors which account for the scores on the subtests, how heavily each of the subtests is weighted with the factors, and which factors appear as a whole for the different groups of subjects. A comparison of factors by sex, grade level, and race was made.

Although there has been a great deal of research in the literature which concerns intelligence and educational achievement of American Indians, no studies were evident in which factor analysis had been applied to achievement test scores for the purpose of comparing full-blooded Indian and white children on large samples. Perhaps the only attempt was made by Harris (33) in which correlations were obtained between 15 language skill variables in four groups of fifth-grade Indian students. The intercorrelations were too irregular for any definite factors to emerge. Experimental studies on intelligence with Indians (20, 26, 27, 28, 29, 30, 34, 36, 39, 46) indicate that differences in test performance between Indian and white children are not necessarily due to differences in innate capacity. The evidence indicates that white children and mixed-blooded Indian children did better where language was vital. The results showed that Indians scored below age and that females performed better than males on the intelligence tests used, but these results seem meager and inconclusive.

Anthropological studies (37, 39, 40, 41) indicated that there were no significant differences in results of performance intelligence tests for Indian children as compared to white children. Intelligence tests of a verbal nature, requiring writing, resulted in lower IQ's for Indian children. The educational achievement of Indians in public schools, federal schools, and mission schools as measured by test scores resulted in a hierarchical order, respectively (2, 13, 42, 43, 52). Indian children with highest degree of Indian blood and non-English speaking ability at the time of entering school obtained the lowest test results. It was believed that cultural factors played an important role in the learning process of these children.

Factor analytic studies on intelligence and age with white children (1, 4, 5, 7, 9, 14, 17, 19, 23, 24, 25, 38, 44, 45, 47, 48) emphasized that a general factor was established early in life. However, conflicting evidence existed as to the first factor

*All footnotes will be found at end of article.

changing with increasing age. Where factor patterns did change, there was no consistent trend in evidence. Factor studies regarding levels and sex for white children (35, 49, 50, 51) revealed a general agreement as to no differentiation in factor patterns between boys and girls. There was indication of conflicting evidence regarding the theory of the general factor playing a less important role as special abilities develop. The factor patterns change relatively little from grade level to grade level, and where changes did occur the patterns did not exhibit any systematic trend. Mental factors seemed to depend on background, education, and experience of the subjects involved.

Where educational skills for white children were factor analyzed (6, 8, 10, 11, 12, 15, 16, 18, 32, 53) it was apparent that educational skills were rather complex with regard to the identification of component factors, that boys and girls vary according to different academic skills, and the more simple and relevant the test the more clearly it produced the factor. Coleman (11) pointed out that a pupil's basic learning difficulty can be diagnosed at an earlier age than was formerly believed and that school achievement depended upon motivation and other factors.

In consideration of the review of literature, five important conclusions were drawn from the results: 1) Factors other than innate capacity were operating on test performance of both Indian and white children; 2) because of contradictory findings, no conclusive statements can be made as to sex and racial superiority on test performance; 3) factor patterns change relatively little from grade to grade with no significant changes between sex; 4) a common factor is established relatively early in life with doubt as to its role becoming minimized as specific abilities develop with increasing age, and 5) because of the complexity of educational skills, it was difficult to identify their component factors.

Hypotheses

On the basis of these studies, one hypothesis concerning differences in achievement raw score means by grade level as to race and sex was developed. Five hypotheses were formulated pertaining to the number, size, and pattern of factors of achievement scores as measured by the California Achievement Test Batteries. The hypotheses are:

1. No differences exist in achievement raw score means on the subtests for the elementary, intermediate, and advanced grade levels for each race by sex.
2. The same number of factors will appear at each grade level for the race groups by sex.
3. The factor loadings will be of the same magnitude for all groups.
4. Higher factor loadings of a verbal nature will appear for white children than for full-blooded Indian children.
5. Higher factor loadings of a numerical nature will appear for white children than for full-blooded Indian children.
6. A common factor will appear throughout all the subtests for all groups.

II. Procedure

The Testing Program

A fall achievement testing program was conducted by the United States Bureau of Indian Affairs and the University of Kansas during the period 1951-55. The California Achievement Test Battery was used throughout the entire testing program for grades four through twelve. The children attending federal schools, public schools, and mission schools were tested in the Albuquerque and Phoenix areas in 1951, Aberdeen area in 1952, Billings area in 1953, and the Anadarko and Muskogee areas in 1954. The tests were administered by Indian Bureau and public and mission school personnel in their respective schools. The public and mission schools included were those which operated in the vicinity of the federal schools included in this study and they enrolled a considerable proportion of Indian pupils. All answer sheets were sent to the University of Kansas for inspection, necessary processing, and scoring. The raw scores of each subtest for each individual were recorded and tabulated. These were punched on IBM cards with background information for each child.

The Population

The geographical locations of the administrative areas of the United States Bureau of Indian Affairs in which testing was conducted were: 1) Phoenix area—Arizona, exclusive of the Navajo and Hopi, 2) Albuquerque area—New Mexico and southern Colorado, exclusive of the Navajo, 3) Aberdeen area—South Dakota, North Dakota, and northeastern Nebraska, 4) Billings area—Montana and Wyoming, 5) Muskogee area—eastern Oklahoma and Mississippi, and 6) Anadarko area—western Oklahoma and Kansas. Between the years 1951 and 1954, inclusive, 13,636 Indian children and 9922 white children were tested in federal, public, and mission schools. These 23,608 children lived in rural areas either on farms, ranches, reservations, or small communities.

The aggregate of full-blooded Indian children attending federal schools and white children attending public schools combined in all areas were chosen for this study. The total number of subjects, 14,888, were separated as to full-blooded Indian children and white children by sex and

TABLE I

DISTRIBUTION OF SUBJECTS BY RACE, SEX, AND GRADE LEVELS

	Full-Blood Indian Children			
	Males		Females	
	N	%	N	%
Elementary Level Grades 4, 5, 6	1223	21.32	1161	20.24
Intermediate Level Grades 7, 8, 9	972	17.81	971	17.79
Advanced Level Grades 10, 11, 12	551	14.91	624	16.89
	White Children			
	Males		Females	
	N	%	N	%
Elementary Level Grades 4, 5, 6	1728	30.13	1624	28.31
Intermediate Level Grades 7, 8, 9	1817	33.30	1697	31.10
Advanced Level Grades 10, 11, 12	1201	32.50	1319	35.70
Total = 14,888				

grade levels. To conform to the suggested grade placement for the achievement tests, the grade levels were combined into elementary (grades 4, 5, and 6), intermediate (grades 7, 8, and 9), and advanced (grades 10, 11, and 12). Table I shows the breakdown of the subjects and the number in each group with percentages. The approximate age ranges for the federal Indian children were from eight to 17 years for the elementary level, 10 to 19 years for the intermediate level, and 15 to 24 years for the advanced level. For public school white children the approximate age ranges were from eight to sixteen years for the elementary level, 11 to 19 years for the intermediate level, and 15 to 20 years for the advanced level.

The Test Battery

The California Achievement Test Battery, Form AA, was utilized to measure the subjects' achievement in fundamental reading, arithmetic, and language skills through the entire testing program. These tests were chosen because they were already being used widely in Indian Bureau schools, they were selected by Indian Bureau personnel as being as close to content of the curriculum as any of the other achievement batteries, and they were machine scored. The tests of basic skills used were: a) reading vocabulary, b) reading comprehension, c) arithmetic reasoning, d) arithmetic fundamentals, e) mechanics of English and grammar, and f) spelling. Since the sections of each test contained such a small number of items, only total scores for each test were obtained rather than scores for each section of the test. The scoring formula was the number of right responses for each of the six tests. Since the California Achievement Test Battery is widely known and is used as a test of academic achievement, a discussion of its construction, reliability, and validity is unnecessary.

The Factorial Analyses

The raw scores for each individual were punched on IBM cards. Tabulating machine operations were used in obtaining frequency distributions, sums of squares, and sums of cross products for each test by race, sex, and grade level groups.[1] The means, medians, standard deviations, ranges, and indices of skewness were computed from the frequency distributions. There were only slight discrepancies with no differences between the median and mean greater than three raw score points. In many cases the differences were less than one. Many of the distributions were slightly skewed which indicated that the tests could have been somewhat too difficult for the students tested. However, the degree of skewness was not sufficiently large where the assumption of normality could not be met. A rather consistent trend was noted in the variability among races, sexes, and grade level groups. In general, the standard deviation is considerably larger on the Reading, Vocabulary, and Arithmetic fundamentals test than on the other four tests in the battery for the various groups. Pearson product-moment correlation coefficients were determined among the six tests by race, sex, and grade level. Twelve correlation matrices for the various groups were formulated, i.e., a separate correlation matrix was constructed for males and females at the elementary, intermediate, and advanced grade level groups for both full-blooded Indian children and white children. The correlations were all positive with the lowest correlation of $r = .35$ between the Reading vocabulary test and the Arithmetic fundamentals test at the advanced level for full-blooded Indian children. For the sample size $N = 551$ this correlation differs from zero at the $.\overline{01}$ level of confidence, as do all other coefficients of correlation in this study.

The correlation matrices were factored by the centroid method of analysis (22, 31, 54). The first extraction of factors was accomplished by using the highest correlation in each matrix as the estimated communality. After each factoring, adjustments were made in the diagonal cells and the process was continued until three factors were extracted. In order to obtain a better estimate of the communalities, the calculated communalities were used in the diagonal cells and the factoring was repeated. The new centroid factor loadings and the revised estimate of the communalities are listed in Tables IV through VII. After the second extraction of factors the residuals of the third factor were sufficiently small to cease factoring.

An oblique simple structure was obtained for the three factors. The method of extended vectors was utilized because, according to Thurstone (54), where three dimensions are used the method gives the solution in one diagram if a simple-structure solution exists. This method seemed most appropriate since a central reference vector existed on which all the test vectors have positive projections. The Extended Factor Matrices E_0 are shown in Table VIII through XIX. Upon oblique rotation of the centroid factor loadings for full-blooded Indian male children at all grade level groups, the third extracted factor rotated with the second did not reveal clear configurations, therefore, meaningful traces could not be drawn which would aid in the establishment and interpretation of these factors. The second factor residuals, range -.05 to .04, for full-blooded Indian male children at all levels are sufficiently small to yield a meaningful third factor. Apparently, the third factor loadings for this group are error residuals. As a result, only factors I and II were used in oblique rotation for these three grade levels in which a single plane, two-dimensional rotation was made. Tables VIII, XIX,

TABLE II

INTERCORRELATIONS BETWEEN THE RAW SCORES OF EACH TEST FOR FULL-BLOOD INDIAN CHILDREN BY SEX ON THE CALIFORNIA ACHIEVEMENT BATTERY

		Full-Blood Indian Children											
		Males						Females					
		(N=1223)						(N=1161)					
Level	Test[a]	1	2	3	4	5	6	1	2	3	4	5	6
Elementary	1		72	72	56	66	64		76	69	54	65	65
	2	72		71	56	69	63	76		70	60	70	66
	3	72	71		73	68	65	69	70		71	65	65
	4	56	56	73		54	48	54	60	71		52	53
	5	66	69	68	54		59	65	70	65	52		64
	6	64	63	65	48	59		65	66	65	53	64	
		(N=972)						(N=971)					
Intermediate	1		68	61	51	64	71		62	54	47	60	63
	2	68		71	63	67	65	62		63	59	64	59
	3	61	71		78	70	56	54	63		71	66	54
	4	51	63	78		65	50	47	59	71		60	48
	5	64	67	70	65		62	60	64	66	60		63
	6	71	65	56	50	62		63	59	54	48	63	
		(N=551)						(N=624)					
Advanced	1		68	53	35	55	63		58	49	37	52	51
	2	68		61	52	60	60	58		61	49	61	57
	3	53	61		70	63	52	49	61		72	59	48
	4	35	52	70		58	45	37	49	72		53	46
	5	55	60	63	58		55	52	61	59	53		59
	6	63	60	52	45	55		51	57	48	46	59	

Note.--Decimals omitted for convenience.

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

TABLE III

INTERCORRELATIONS BETWEEN THE RAW SCORES OF EACH TEST FOR WHITE CHILDREN BY SEX ON THE CALIFORNIA ACHIEVEMENT BATTERY

		White Children											
		Males						Females					
Level	Test[a]	1	2	3	4	5	6	1	2	3	4	5	6
				(N=1728)						(N=1624)			
Elementary	1		76	75	55	67	64		77	75	59	67	67
	2	76		73	57	67	65	77		76	63	66	65
	3	75	73		73	68	62	75	76		75	65	63
	4	55	57	73		56	51	59	63	75		54	53
	5	67	67	68	56		61	67	66	65	54		61
	6	64	65	62	51	61		67	65	63	53	61	
				(N=1817)						(N=1697)			
Intermediate	1		76	67	54	65	67		74	66	59	67	62
	2	76		73	63	68	61	74		71	65	68	56
	3	67	73		75	68	57	66	71		77	68	54
	4	54	63	75		64	53	59	65	77		67	52
	5	65	68	68	64		61	67	68	68	67		58
	6	67	61	57	53	61		62	56	54	52	58	
				(N=1201)						(N=1319)			
Advanced	1		72	63	49	58	61		71	59	46	61	62
	2	72		63	54	58	53	71		62	53	61	56
	3	63	63		75	67	53	59	62		76	63	48
	4	49	54	75		58	43	46	53	76		59	46
	5	58	58	67	58		60	61	61	63	59		64
	6	61	53	53	43	60		62	56	48	46	64	

Note.--Decimals omitted for convenience.

[a] Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

TABLE IV

CENTROID FACTOR MATRICES AND OBTAINED COMMUNALITIES WITH PROPORTIONS OF VARIANCES CONTRIBUTED BY THE CENTROID FACTORS (F_o)
(MALE)

		Full-Blood Indian Children						
		Male						
		Factor Loadings			Factor Variances			h^2
Level	Test[a]	I	II	III	I'	II'	III'	
Elementary	1	832	123	036	692	015	001	708
	2	840	162	155	706	026	024	756
	3	888	-192	-131	789	037	017	843
	4	737	-438	131	543	192	017	752
	5	788	100	096	621	010	009	640
	6	747	285	-287	558	081	082	721
Intermediate	1	798	324	-043	637	105	002	744
	2	833	079	149	694	006	022	722
	3	861	-281	064	741	079	004	824
	4	787	-353	043	619	124	002	745
	5	814	-036	-064	663	001	004	668
	6	769	317	-107	591	100	011	702
Advanced	1	758	425	151	575	181	023	779
	2	805	184	-060	648	034	004	686
	3	821	-272	151	674	074	023	771
	4	723	-425	-166	523	181	028	732
	5	766	-070	-060	587	005	004	596
	6	723	203	-075	523	041	006	570

Note.--Decimals omitted for convenience.

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

TABLE V

CENTROID FACTOR MATRICES AND OBTAINED COMMUNALITIES WITH PROPORTIONS OF VARIANCES CONTRIBUTED BY THE CENTROID FACTORS (F_O)
(FEMALE)

		Full-Blood Indian Children						
		Female						
		Factor Loadings			Factor Variances			h^2
Level	Test[a]	I	II	III	I'	II'	III'	
Elementary	1	830	175	104	689	031	011	731
	2	863	140	-070	745	020	005	770
	3	859	-236	104	738	056	011	805
	4	735	-315	017	540	099	000	639
	5	793	166	-122	629	028	015	672
	6	776	114	157	602	013	025	640
	1	750	284	078	563	081	006	650
	2	795	041	098	632	002	010	644
	3	816	-276	-039	666	076	002	744
	4	756	-333	020	572	111	000	683
	5	814	-032	-157	663	001	025	689
	6	752	268	-118	566	072	014	652
	1	675	258	-167	456	067	028	551
	2	782	176	-143	612	031	020	663
	3	820	-326	-084	672	106	007	785
	4	729	-360	143	531	130	020	681
	5	769	109	108	591	012	012	615
	6	715	244	191	511	060	036	607

Note.--Decimals omitted for convenience.

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language 6. Spelling

TABLE VI

CENTROID FACTOR MATRICES AND OBTAINED COMMUNALITIES WITH PROPORTIONS OF VARIANCES CONTRIBUTED BY THE CENTROID FACTORS (F_O)
(MALE)

		White Children						
		Male						
		Factor Loadings			Factor Variances			h^2
Level	Test[a]	I	II	III	I'	II'	III'	
Elementary	1	859	176	170	738	031	029	798
	2	850	166	099	723	028	010	761
	3	887	-213	141	787	045	020	852
	4	739	-324	-085	546	105	007	658
	5	787	065	-099	619	004	010	633
	6	748	139	-113	560	019	013	592
Intermediate	1	837	334	093	701	111	009	821
	2	864	109	172	746	012	030	788
	3	856	-226	159	733	051	025	809
	4	784	-360	-026	615	130	001	746
	5	804	-025	-106	646	001	011	658
	6	739	142	-199	546	020	040	606
Advanced	1	806	341	207	650	116	043	809
	2	787	203	197	619	041	039	699
	3	855	-283	128	731	080	016	827
	4	750	-334	089	563	112	008	683
	5	780	051	-197	608	003	039	650
	6	699	167	-197	489	028	039	556

Note.--Decimals omitted for convenience.

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling.

TABLE VII

CENTROID FACTOR MATRICES AND OBTAINED COMMUNALITIES WITH PROPORTIONS OF VARIANCES CONTRIBUTED BY THE CENTROID FACTORS (F_o)
(FEMALE)

Level	Test[a]	White Children						
		Female						
		Factor Loadings			Factor Variances			h^2
		I	II	III	I'	II'	III'	
Elementary	1	864	181	-101	746	033	010	789
	2	866	078	-068	750	006	005	761
	3	891	-207	051	794	043	003	840
	4	764	-294	135	584	086	018	688
	5	766	233	220	587	054	084	725
	6	756	164	-017	572	027	000	599
Intermediate	1	834	274	122	696	075	015	786
	2	844	091	191	712	008	036	756
	3	854	-246	109	729	061	012	802
	4	815	-301	-054	664	091	003	758
	5	811	009	-068	658	000	005	663
	6	698	173	-191	487	030	036	553
Advanced	1	796	291	-226	634	045	051	730
	2	798	132	-129	637	017	017	671
	3	824	-357	-097	679	127	009	815
	4	754	-390	129	567	152	017	736
	5	796	072	194	634	005	038	677
	6	728	271	151	530	073	023	626

Note.--Decimals omitted for convenience.

[a] Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

TABLE VIII

MATRICES OF OBLIQUE ROTATION OF AXES FOR FULL-BLOOD INDIAN CHILDREN, MALES, ELEMENTARY GRADE LEVEL GROUP

Unrotated Factor Matrix F_o

Test[a]	I	II
1	.832	.123
2	.840	.162
3	.888	-.192
4	.737	-.438
5	.788	.100
6	.747	.285

Rotated Oblique Factor Matrix V

Test[a]	A	B
1	.424	-.051
2	.463	-.014
3	.152	-.371
4	-.133	-.581
5	.386	-.064
6	.543	-.125

Matrix S = L. Direction Coefficients of a Normal to Each Coordinate Plane

Factor	A'	B'
I	.40	-.21
II	1.00	1.00

Matrix Λ. Direction Cosines of Reference Vectors Λp

Factor	A	B
I	.372	-.206
II	.929	.979

Matrix $C = \Lambda'\Lambda$. Cosines of Angles Between Reference Vectors Λpq

Factor	A	B
A	1.000	
B	.833	1.000

Extended Factor Matrix E_o

Test[a]	De	I	II	III
1	1.20192	1.000	.148	.043
2	1.19048	1.000	.193	.185
3	1.12613	1.000	-.216	-.148
4	1.35685	1.000	-.594	.178
5	1.26904	1.000	.127	.122
6	1.33869	1.000	.382	-.384

[a]Test:--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

TABLE IX

MATRICES OF OBLIQUE ROTATION OF AXES FOR FULL-BLOOD INDIAN CHILDREN, MALES, INTERMEDIATE GRADE LEVEL GROUP

Unrotated Factor Matrix F_o

Test[a]	I	II
1	.798	.324
2	.833	.079
3	.861	-.281
4	.787	-.353
5	.814	-.036
6	.769	.317

Rotated Oblique Factor Matrix V

Test[a]	A	B
1	.521	.081
2	.294	-.163
3	-.047	-.516
4	-.136	-.564
5	.178	-.268
6	.507	.083

Matrix S = L. Direction Coefficients of a Normal to Each Coordinate Plane

Factor	A'	B'
I	.27	-.30
II	1.00	1.00

Matrix Λ. Direction Cosines of Reference Vectors Λ_p

Factor	A	B
I	.261	-.287
II	.966	.958

Matrix $C = \Lambda'\Lambda$. Cosines of Angles Between Reference Vectors Λ_{pq}

Factor	A	B
A	1.000	
B	.851	1.000

Extended Factor Matrix E_o

Test[a]	De	I	II	III
1	1.25313	1.000	.406	-.054
2	1.20048	1.000	.095	.179
3	1.16144	1.000	-.326	.074
4	1.27065	1.000	-.449	.055
5	1.22850	1.000	-.044	-.079
6	1.30039	1.000	.412	-.139

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

TABLE X

MATRICES OF OBLIQUE ROTATION OF AXES FOR FULL-BLOOD INDIAN CHILDREN, MALES, ADVANCED GRADE LEVEL GROUP

Unrotated Factor Matrix F_o

Test[a]	I	II
1	.758	.425
2	.805	.184
3	.821	-.272
4	.723	-.425
5	.766	-.070
6	.723	.203

Rotated Oblique Factor Matrix V

Test[a]	A	B
1	.647	.143
2	.433	-.010
3	.007	-.534
4	-.170	-.645
5	.180	-.326
6	.041	-.054

Matrix S = L. Direction Coefficients of a Normal to Each Coordinate Plane

Factor	A'	B'
I	.34	-.36
II	1.00	1.00

Matrix Λ . Direction Cosines of Reference Vectors Λ p

Factor	A	B
I	.322	-.339
II	.947	.941

Matrix C = $\Lambda'\Lambda$. Cosines of Angles Between Reference Vectors Λ pq

	A	B
A	1.000	
B	.785	1.000

Extended Factor Matrix E_o

Test[a]	De	I	II	III
1	1.31926	1.000	.351	.199
2	1.24224	1.000	.229	-.075
3	1.21803	1.000	-.331	.184
4	1.38313	1.000	-.588	-.230
5	1.30548	1.000	-.091	-.078
6	1.38313	1.000	.281	-.104

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals. 5. Language, 6. Spelling

TABLE XI

MATRICES OF OBLIQUE ROTATION OF AXES FOR FULL-BLOOD INDIAN CHILDREN, FEMALES, ELEMENTARY GRADE LEVEL GROUP

Extended Factor Matrix E_o

Test[a]	D_e	I	II	III
1	1.20482	1.000	.211	.125
2	1.15875	1.000	.162	-.081
3	1.16414	1.000	-.275	.121
4	1.36054	1.000	-.429	.023
5	1.26103	1.000	.209	-.154
6	1.28866	1.000	.147	.202

Rotated Oblique Factor Matrix V

Test[a]	A	B	C
1	.189	-.019	.043
2	.014	.025	.211
3	.050	.390	-.025
4	-.063	.446	.029
5	-.045	-.012	.257
6	.216	.030	-.027

Matrix S = L. Direction Coefficients of a Normal to Each Coordinate Plane

Factor	A'	B'	C'
I	.04	.19	.14
II	.36	-1.00	.18
III	1.00	- .02	-1.00

Matrix Λ. Direction Cosines of Reference Vectors Λ_p

Factor	A	B	C
I	.038	.187	.137
II	.339	-.982	.176
III	.940	-.020	-.975

Matrix C = $\Lambda'\Lambda$. Cosines of Angles Between Reference Vectors Λ_{pq}

Factor	A	B	C
A	1.000	-.345	-.852
B	- .345	1.000	-.128
C	- .852	-.128	1.000

Matrix T. Direction Cosines of Primary Vectors T_p

Factor	A	B	C
I	.969	.946	.980
II	.181	-.316	.172
III	.168	.076	-.098

Matrix TT' = Rpq. Cosines of Angles Between Primary Vectors T_p[b]

Factor	A	B	C
A	1.000		
B	.872	1.001	
C	.964	.865	1.000

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

[b]Factor A should read Factor C
Factor C should read Factor A

TABLE XII

MATRICES OF OBLIQUE ROTATION OF AXES FOR FULL-BLOOD INDIAN CHILDREN, FEMALES, INTERMEDIATE GRADE LEVEL GROUP

Extended Factor Matrix E_o

Test[a]	D_e	I	II	III
1	1.33333	1.000	.379	.104
2	1.25786	1.000	.052	.123
3	1.22549	1.000	-.338	-.048
4	1.32275	1.000	-.440	.026
5	1.22850	1.000	-.039	-.193
6	1.32979	1.000	.356	-.157

Rotated Oblique Factor Matrix V

Test[a]	A	B	C
1	.232	-.015	.048
2	.196	.228	-.017
3	.075	.537	.053
4	.023	.569	-.021
5	-.065	.310	.218
6	.040	.006	.236

Matrix S = L. Direction Coefficients of a Normal to Each Coordinate Plane

Factor	A'	B'	C'
I	.12	.36	.09
II	.26	-1.00	.21
III	1.00	- .03	-1.00

Matrix Λ . Direction Cosines of Reference Vectors Λ_p

Factor	A	B	C
I	.115	.339	.088
II	.250	-.941	.205
III	.961	-.028	-.975

Matrix C = $\Lambda'\Lambda$. Cosines of Angles Between Reference Vectors Λ_{pq}

Factor	A	B	C
A	.999		
B	-.223	1.001	
C	-.876	- .136	1.000

Matrix T. Direction Cosines of Primary Vectors T_p

Factor	A	B	C
I	.931	.913	.921
II	.331	-.408	.337
III	.153	-.004	-.197

Matrix TT' = Rpq. Cosines of Angles Between Primary Vectors T_p

Factor	A	B	C
A	1.000		
B	.714	1.000	
C	.939	.704	1.001

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

TABLE XIII

MATRICES OF OBLIQUE ROTATION OF AXES FOR FULL-BLOOD INDIAN CHILDREN, FEMALES, ADVANCED GRADE LEVEL GROUP

Extended Factor Matrix E_o

Test[a]	D_e	I	II	III
1	1.48148	1.000	.382	-.247
2	1.27877	1.000	.225	-.183
3	1.21951	1.000	-.398	-.102
4	1.37174	1.000	-.494	.196
5	1.30039	1.000	.142	.140
6	1.39860	1.000	.341	.267

Rotated Oblique Factor Matrix V

Test[a]	A	B	C
1	-.009	-.030	.338
2	-.003	.075	.317
3	-.109	.554	.161
4	.085	.530	-.081
5	.210	.101	.059
6	.327	-.053	-.002

Matrix S = L. Direction Coefficients of a Normal to Each Coordinate Plane

Factor	A'	B'	C'
I	.10	.30	.19
II	.35	-1.00	.22
III	1.00	- .14	-1.00

Matrix Λ . Direction Cosines of Reference Vectors Λp

Factor	A	B	C
I	.094	.285	.182
II	.329	-.949	.211
III	.940	-.133	-.960

Matrix $C = \Lambda'\Lambda$. Cosines of Angles Between Reference Vectors Λpq

Factor	A	B	C
A	1.000		
B	-.410	1.000	
C	-.817	-.031	.999

Matrix T. Direction Cosines of Primary Vectors T_p

Factor	A	B	C
I	.940	.890	.930
II	.249	-.452	.307
III	.233	.069	-.201

Matrix $TT' = Rpq$. Cosines of Angles Between Primary Vectors T_p[b]

Factor	A	B	C
A	1.000		
B	.740	1.001	
C	.904	.675	1.000

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

[b]Factor A should read Factor C
Factor C should read Factor A

TABLE XIV

MATRICES OF OBLIQUE ROTATION OF AXES FOR WHITE CHILDREN, MALES, ELEMENTARY GRADE LEVEL GROUP

Extended Factor Matrix E_o

Test[a]	D_e	I	II	III
1	1.16414	1.000	.204	.198
2	1.17647	1.000	.195	.116
3	1.12740	1.000	-.240	.159
4	1.35318	1.000	-.438	-.115
5	1.27065	1.000	.083	-.126
6	1.33690	1.000	.186	-.151

Rotated Oblique Factor Matrix V

Test[a]	A	B	C
1	.299	-.001	-.024
2	.229	-.005	.042
3	.144	.377	-.064
4	-.119	.423	-.049
5	.003	.050	.208
6	.011	-.030	.231

Matrix S = L. Direction Coefficients of a Normal to Each Coordinate Plane

Factor	A'	B'	C'
I	.10	.17	.13
II	.36	-1.00	.19
III	1.00	.17	-1.00

Matrix Λ. Direction Cosines of Reference Vectors Λp

Factor	A	B	C
I	.094	.165	.127
II	.337	-.972	.185
III	.937	.165	-.975

Matrix $C = \Lambda'\Lambda$. Cosines of Angles Between Reference Vectors Λpq

Factor	A	B	C
A	1.000		
B	- .157	.999	
C	- .839	-.320	1.000

Matrix T. Direction Cosines of Primary Vectors T_p

Factor	A	B	C
I	.958	.921	.979
II	.192	-.387	.141
III	.162	.046	-.149

Matrix $TT' = Rpq$. Cosines of Angles Between Primary Vectors T_p

Factor	A	B	C
A	1.000		
B	.825	1.000	
C	.951	.840	1.001

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

TABLE XV

MATRICES OF OBLIQUE ROTATION OF AXES FOR WHITE CHILDREN, MALES, INTERMEDIATE GRADE LEVEL GROUP

Extended Factor Matrix E_o

Test[a]	D_e	I	II	III
1	1.19474	1.000	.399	.111
2	1.15741	1.000	.126	.199
3	1.16822	1.000	-.264	.186
4	1.27551	1.000	-.459	-.033
5	1.24378	1.000	-.031	-.132
6	1.35318	1.000	.192	-.269

Rotated Oblique Factor Matrix V

Test[a]	A	B	C
1	.372	-.134	.045
2	.325	.106	-.050
3	.138	.406	-.068
4	-.098	.450	.094
5	.013	.117	.205
6	.012	-.081	.304

Matrix S = L. Direction Coefficients of a Normal to Each Coordinate Plane

Factor	A'	B'	C'
I	.17	.18	.13
II	.62	-1.00	.09
III	1.00	.40	-1.00

Matrix Λ. Direction Cosines of Reference Vectors Λp

Factor	A	B	C
I	.143	.165	.128
II	.522	-.916	.089
III	.841	.366	-.988

Matrix $C = \Lambda'\Lambda$. Cosines of Angles Between Reference Vectors Λpq

Factor	A	B	C
A	1.000		
B	- .147	1.000	
C	- .766	-.422	1.000

Matrix T. Direction Cosines of Primary Vectors T_p

Factor	A	B	C
I	.962	.919	.972
II	.232	-.387	.087
III	.164	.084	-.219

Matrix $TT' = Rpq$. Cosines of Angles Between Primary Vectors T_p

Factor	A	B	C
A	1.001		
B	.807	1.001	
C	.923	.841	1.000

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

TABLE XVI

MATRICES OF OBLIQUE ROTATION OF AXES FOR WHITE CHILDREN, MALES, ADVANCED GRADE LEVEL GROUP

Extended Factor Matrix E_o

Test[a]	D_E	I	II	III
1	1.24070	1.000	.423	.257
2	1.27065	1.000	.258	.250
3	1.16959	1.000	-.331	.150
4	1.33333	1.000	-.445	.119
5	1.28205	1.000	.065	-.252
6	1.43062	1.000	.239	-.282

Rotated Oblique Factor Matrix V

Test[a]	A	B	C
1	.463	-.068	.015
2	.372	.053	-.003
3	.037	.499	-.008
4	-.036	.510	.001
5	.033	.075	.350
6	.025	-.051	.354

Matrix S = L. Direction Coefficients of a Normal to Each Coordinate Plane

Factor	A'	B'	C'
I	.15	.25	.20
II	.74	-1.00	.18
III	1.00	.32	-1.00

Matrix Λ. Direction Cosines of Reference Vectors Λp

Factor	A	B	C
I	.120	.232	.193
II	.591	-.927	.174
III	.798	.296	-.966

Matrix C = $\Lambda'\Lambda$. Cosines of Angles Between Reference Vectors Λpq

Factor	A	B	C
A	1.000		
B	- .284	1.000	
C	- .645	-.402	1.000

Matrix T. Direction Cosines of Primary Vectors T_p

Factor	A	B	C
I	.921	.928	.953
II	.307	-.353	.156
III	.239	.121	-.258

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

Matrix $TT' = Rpq$. Cosines of Angles Between Primary Vectors T_p

Factor	A	B	C
A	1.000		
B	.775	1.000	
C	.864	.798	.999

TABLE XVII

MATRICES OF OBLIQUE ROTATION OF AXES FOR WHITE CHILDREN, FEMALES, ELEMENTARY GRADE LEVEL GROUP

Extended Factor Matrix E_o

Test[a]	D_a	I	II	III
1	1.15741	1.000	.209	-.117
2	1.15473	1.000	.090	-.079
3	1.12233	1.000	-.232	.057
4	1.30890	1.000	-.385	.177
5	1.30548	1.000	.304	.287
6	1.32275	1.000	.217	-.022

Rotated Oblique Factor Matrix V

Test[a]	A	B	C
1	.189	.030	.298
2	-.026	-.076	.247
3	-.021	.385	.082
4	.023	.467	-.040
5	.298	-.004	-.020
6	.052	-.012	.193

Matrix S = L. Direction Coefficients of a Normal to Each Coordinate Plane

Factor	A'	B'	C'
I	.01	.21	.20
II	.40	-1.00	.20
III	1.00	.32	-1.00

Matrix Λ . Direction Cosines of Reference Vectors Λp

Factor	A	B	C
I	.009	.196	.192
II	.371	-.934	.192
III	.929	.299	-.962

Matrix $C = \Lambda'\Lambda$. Cosines of Angles Between Reference Vectors Λpq

Factor	A	B	C
A	1.000		
B	- .067	1.000	
C	- .821	-.429	.999

Matrix T. Direction Cosines of Primary Vectors T_p

Factor	A	B	C
I	.932	.937	.981
II	.273	-.327	.179
III	.240	.122	-.081

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

Matrix $TT' = Rpq$. Cosines of Angles Between Primary Vectors T_p[b]

Factor	A	B	C
A	1.001		
B	.813	1.000	
C	.944	.851	1.001

[b]Factor A should read Factor C
Factor C should read Factor A

TABLE XVIII

MATRICES OF OBLIQUE ROTATION OF AXES FOR WHITE CHILDREN, FEMALES, INTERMEDIATE GRADE LEVEL GROUP

Extended Factor Matrix E_o

Test[a]	D_e	I	II	III
1	1.19904	1.000	.329	.146
2	1.18483	1.000	.108	.226
3	1.17096	1.000	-.288	.128
4	1.22699	1.000	-.369	-.066
5	1.23305	1.000	.011	-.084
6	1.43267	1.000	.248	-.274

Rotated Oblique Factor Matrix V

Test[a]	A	B	C
1	.325	-.094	.083
2	.308	.086	-.034
3	.089	.413	-.054
4	-.085	.435	.079
5	.037	.131	.182
6	-.014	-.067	.330

Matrix S = L. Direction Coefficients of a Normal to Each Coordinate Plane

Factor	A'	B'	C'
I	.13	.19	.15
II	.49	-1.00	.31
III	1.00	.15	-1.00

Matrix Λ . Direction Cosines of Reference Vectors Λp

Factor	A	B	C
I	.116	.185	.142
II	.437	-.972	.293
III	.892	.146	-.946

Matrix C = $\Lambda'\Lambda$. Cosines of Angles Between Reference Vectors Λpq

Factor	A	B	C
A	1.000		
B	- .183	1.000	
C	- .699	- .397	1.000

Matrix T. Direction Cosines of Primary Vectors T_p

Factor	A	B	C
I	.954	.943	.968
II	.213	-.330	.154
III	.209	.039	-.202

Matrix TT' = Rpq. Cosines of Angles Between Primary Vectors T_p

Factor	A	B	C
A	.999		
B	.837	1.000	
C	.914	.854	1.001

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

TABLE XIX

MATRICES OF OBLIQUE ROTATION OF AXES FOR WHITE CHILDREN, FEMALES, ADVANCED GRADE LEVEL GROUP

Extended Factor Matrix E_o

Test[a]	D_e	I	II	III
1	1.25628	1.000	.366	-.284
2	1.25313	1.000	.165	-.162
3	1.21359	1.000	-.433	-.118
4	1.32626	1.000	-.517	.171
5	1.25628	1.000	.090	.244
6	1.37363	1.000	.372	.207

Rotated Oblique Factor Matrix V

Test[a]	A	B	C
1	-.030	-.079	.425
2	.011	.075	.293
3	-.112	.554	.146
4	.082	.566	-.090
5	.296	.129	-.030
6	.311	-.080	.048

Matrix S = L. Direction Coefficients of a Normal to Each Coordinate Plane

Factor	A'	B'	C'
I	.12	.26	.18
II	.34	-1.00	.26
III	1.00	- .01	-1.00

Matrix Λ. Direction Cosines of Reference Vectors Λp

Factor	A	B	C
I	.113	.252	.172
II	.320	-.968	.248
III	.941	-.010	-.954

Matrix C = $\Lambda'\Lambda$. Cosines of Angles Between Reference Vectors Λpq

Factor	A	B	C
A	1.000		
B	- .291	1.000	
C	- .799	- .187	1.001

Matrix T. Direction Cosines of Primary Vectors T_p

Factor	A	B	C
I	.942	.893	.948
II	.243	-.447	.248
III	.233	.045	-.199

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

[b]Factor A should read Factor C
Factor C should read Factor A

Matrix TT' = Rpq. Cosines of Angles Between Primary Vectors T_p[b]

Factor	A	B	C
A	1.001		
B	.743	.999	
C	.907	.727	1.000

TABLE XX

FACTOR PATTERNS AND FACTOR STRUCTURES BY RACE, SEX, AND GRADE LEVEL GROUPS AFTER PERMUTATED ROTATION

Full-Blood Indian Children

	Males									Females								
	Elementary			Intermediate			Advanced			Elementary			Intermediate			Advanced		
Factor	A	B	C	A	B	C	A	B	C	A[b]	B	C[b]	A	B	C	A[b]	B	C[b]
Test[a]																		
1	.42			.52			.65						.23			.34		
2	.46			.29			.43			.21			.20	.23		.32		
3		-.37			-.52			-.53			.39			.54			.49	
4		-.58			-.56			-.65			.45			.57			.47	
5	.39				-.27			-.33		.26				.31	.22			.21
6	.54			.51								.22			.24			.33
	White Children																	
1	.30			.37			.46			.30			.33			.43		
2	.23			.33			.37			.25			.31			.29		
3		.38			.41			.50			.39			.41			.55	
4		.42			.45			.51			.47			.44			.57	
5			.21			.21			.35			.30						.30
6			.23			.30			.35						.33			.31

[a]Test.--1. Reading Vocabulary, 2. Reading Comprehension, 3. Arithmetic Reasoning, 4. Arithmetic Fundamentals, 5. Language, 6. Spelling

[b]Factor A should read Factor C
Factor C should read Factor A

and X give the unrotated factor loadings, the rotated oblique factor loadings, the direction coefficients of the normal to each co-ordinate plane, the transformation matrix, and the matrix of cosines of the angular separations of the oblique axes for full-blooded Indian children at the elementary, intermediate, and advanced levels, respectively.

The direction coefficients of a normal to each co-ordinate plane, Matrix S, when normalized, resulted in the transformation Matrix Λ. The column numbers represent the direction cosines of the reference vectors with respect to the unrotated factor Matrix F_0. Post-multiplying the unrotated factor matrix by the transformation matrix yielded the rotated oblique factor Matrix V. The cosines of angles between the reference vectors, Matrix C (intercorrelations)' were obtained by multiplying the matrix of direction cosines of reference vectors by its transformation.

After the oblique rotations were accomplished for all the other groups, the cosines of angles between the reference vectors all had negative signs indicating a positive correlation between factors A and B, factors A and C, and factors B and C. The cosines between the reference vectors are shown in Matrix C of Tables VIII through XIX, and the cosines between the primary factors are shown in Matrix R_{pq} of Tables XI through XIX. The intercorrelations were computed by using

$$T = D\Lambda^{-1} \quad \text{and} \quad R_{pq} = D\Lambda^{-1}(D\Lambda^{-1})' = TT'$$

where T is the inverse of Matrix Λ normalized or the factor matrix for the primary vectors and R_{pq} is the correlation between the primary factors which is obtained by matrix multiplication of Matrix T by its transpose. The rotated oblique factor Matrix V for these groups was determined by post-multiplying the unrotated factor Matrix F_0 of Tables V, VI, and VII by the respective transformation Matrix Λ of Tables XI through XIX.

Matrix C in Tables VIII through XIX show the intercorrelations between the reference vectors while Matrix R_{pq} in Tables XI through XIX show intercorrelations between the primary vectors. These correlations were all significantly different from zero. This indicates that the obtained factors were not independent and that a relationship existed among them. The highest correlation existed between factors A and C which had significant loadings on Tests 1, 2, 5, and 6. In all cases, Tests 3 and 4 had significant loadings of Factor B which was also significantly related to the other two factors. These relationships indicated the presence of a second-order factor common to all the tests.

In order to facilitate comparison of the various groups, factor patterns of the rotated oblique factor loadings on each of the six tests are presented in Table XX. Factor loadings of less than .20 were not included in the pattern plan. The elementary and advanced grade level groups for females of both races showed patterns in a reverse direction. Permutated rotation was used to reverse the pattern so that comparisons of the sets of factor patterns could be made more readily (55). Factors A and C in Matrix V in Tables XI, XIII, XVII, and XIX, for these two grade level groups, were interchanged by post-multiplying by the permutation matrix. Correspondingly, the interchange in Matrix R_{pq} was obtained by pre-multiplying and post-multiplying it by the permutation matrix which resulted in an interchange of r_{AB} and r_{CB}. This reversal did not effect the values of the factor loadings.

Two significant factors were present for full-blooded Indian male children at each grade level, while three significant factors were present for the other groups.

Rather consistent patterns of factors were evident among grade level groups for full-blooded Indian male children on all tests except 5 and 6. Test 5 changed from a Factor A to a Factor B loading at the intermediate and advanced levels, while Factor A vanished on Test 6 at the advanced level. For full-blooded Indian female children a rather similar pattern existed among the three grade level groups. The "staircase effect" was present at all levels with the elementary group showing a Factor C loading on Test 5 and no significant loading on Test 1. A similar pattern was noted at the intermediate level with Factor B loadings on Tests 2 and 5. For white male and female children at all grade levels, three groups of two tests each, that is, Tests 1 and 2, 3 and 4, and 5 and 6 had significant loadings except for females on Test 5 at the intermediate level and on Test 6 at the elementary level.

The largest factor loadings were evident for full-blooded Indian male children. In most cases the loadings tended to increase by grade level groups for each race and sex. The sets of factor patterns between race groups appeared to be more varied than between sex. This was noted especially for females where the factors clustered in groups of two on the same tests at the various levels for white children. The same pattern was noted at the advanced level for full-blooded Indian children, but at the other two levels greater pattern variability was present. In relation to sex, the factor patterns were similar for white children, but they showed more inconsistency for the full-blooded Indian children.

III. Interpretation

Since the tests used in this study were all paper-and-pencil variety, and since the tests purported to measure the basic educational skills, the nature of the factors were considered in view of educational achievement. This limitation was un-

desirable, but restrictive because no other measurements were available for the subjects studied.

In view of the tests' content, Factors A and C were identified as verbal factors and Factor B was identified as a numerical factor with a common factor underlying all the tests, which might be identified as a Schooling factor. It would be inappropriate to attempt an interpretation of the abilities necessary to perform the tasks called for by the various tests. The tests with more verbal material, Reading vocabulary, Reading comprehension, Language, and Spelling did not remain clustered as well as did those tests with number material, Arithmetic reasoning, and Arithmetic fundamentals.

Evaluation of Hypotheses[2]

Hypothesis 1—A comparison between the mean raw scores computed by the critical ratio showed a significant difference, .05 level of confidence, on all tests for the three grade level groups in favor of white children. With regard to sex differences, the mean raw scores for female white children were significantly higher, .05 level of confidence, than for male white children with one exception: At the advanced grade level males surpassed females on the Arithmetic reasoning test. In the majority of cases, the mean raw scores were significantly higher, .05 level of confidence, for full-blooded Indian female children. The mean differences in favor of full-blooded Indian male children were on the Reading comprehension and Arithmetic reasoning tests at the elementary level, Reading vocabulary, Reading comprehension, and Arithmetic fundamentals at the intermediate level, and Reading vocabulary and Reading comprehension at the advanced level.

Hypothesis 2—Two factors were found for full-blooded Indian male children at the elementary, intermediate, and advanced grade levels. These were identified as a verbal and a numerical factor. The high correlations between these factors suggested a second-order general factor. Three factors were found for full-blooded Indian female children and for male and female white children at each of the three grade level groups.

Hypothesis 3—The factor loadings ranged from .20 to .65 for full-blooded Indian children and from .21 to .57 for white children. More high loadings appeared for full-blooded Indian male children than for any other race-sex group.

Hypothesis 4—The factor loadings on the reading tests were higher for full-blooded Indian male children than for white male children. In general, the differences in loadings were .10 or greater in favor of the full-blooded Indian children. A similar difference was evident on the language and spelling tests.

Hypothesis 5—Full-blooded Indian male children showed higher loadings of a number factor on the arithmetic tests than did white male children. Generally, the size of the loadings of this factor increased with increase in grade level. For full-blooded Indian female children the loadings were highest at the intermediate level and lowest at the elementary level, whereas, for the white female children the loadings were highest at the advanced level and approximately of the same magnitude at the elementary and intermediate level with moderate loadings.

Hypothesis 6—The correlation matrices for all groups analyzed showed highly significant positive relationships. This suggested an underlying common factor. The common factor was substantiated by the significant intercorrelations existing between the reference and the primary vectors.

IV. Discussion

There is little evidence regarding differences in sets of factor patterns and factor structures by race and sex for the elementary, intermediate, and advanced grade level groups. Although the results did not conform ideally to Garrett's (25) differential hypothesis whereby abstract intelligence changes in organization from a general ability to an organized group of abilities or factors as age increases, the arithmetic tests showed rather consistent increases in factor loadings between the grade levels from the elementary to the advanced level. A similar trend was noted for both full-blooded Indian and white children on the reading tests which was more pronounced for males. This evidence was based on cross-sectional rather than longitudinal data. When the correlations between the reference vectors and the primary vectors are considered, in most cases the correlations decreased with increase in grade level. These correlations were significantly different from each other which indicated differentiation between grade level groups.

With regard to the interpretation of the factors it should be pointed out that with only six variables the reference frames would not be too stable, therefore, the factor loadings might be subject to change from sample to sample. It was difficult to adequately establish the traces on the test projections with only six tests and criterion tests with established factors for the subjects in this study were not available.

The high positive correlations among the tests and among the primary vectors strongly suggested the presence of a factor in the second-order domain. Apparently, this accounted for the low loadings on the first-order factors. It may be assumed that the tests are rather complex in that they not only involve a specific factor, but also the common factor defined by the tests. Tests 1, 2, 5, and 6 appeared to be more complex than Tests 3 and 4. The reading and language tests involved two factors in addition to the common fac-

tor. For full-blooded Indian male children the language test clustered with the arithmetic tests at the intermediate and advanced levels. Perhaps at these levels the three tests have some associative characteristics. A similar trend was noted for full-blooded Indian female children at the intermediate level with the reading comprehension test also joining the cluster. The clustering of the reading comprehension and language tests with the arithmetic tests did not occur for the other groups.

Since full-blooded Indian children and white children receive their education in American schools, it seems reasonable to assume that the Indian children should learn the same basic skills as the white children. The results as shown in this study as well as in other studies cited herein, indicated that significant differences existed between races in educational achievement. Apparently, factors other than school experiences play an important role. These factors could be genetic, environmental, and experiential. According to French (21), the factors might readily be conceptualized as experiential factors since they were obtained from measures on achievement tests. French (21) named the common factor which involves skills acquired in school, e.g., fractions, decimals, arithmetic problems, and vocabulary tests, as Schooling. This factor usually appears in the second-order domain. He calls attention to the interpretation of the general, Number, and Verbal factors which are all quite distinct, and the relation between the Verbal or Number factor has nothing to do with the actual mechanism of Verbal or Number ability. It is perhaps an arbitrary requirement of American schools which is determined by differences in amount of formal education. The findings in this study seem to corroborate French's interpretation.

Since the factor patterns and size of factor loadings are essentially the same for both racial groups by grade level and sex, it indicates that a large proportion of the content of the California Achievement Test Battery measures approximately the same basic educational skills at each grade level. This evidence of continuity of content and the variability of mean achievement suggests that various other factors are interferring with good articulation between elementary and intermediate and between intermediate and advanced grade levels for the two racial groups.

The significant differences in mean achievement and the slight discrepancies in the sets of factor patterns between full-blooded Indian and white children might be speculated from a theoretical viewpoint. The groups studied were reasonably comparable with respect to geographical location and educational grade placement. Assuming that these conditions were sufficiently controlled, it seems that other factors would account for the differences obtained. American culture perhaps determines very largely that both Indian children and white children develop basic educational skills since these skills are important to both races outside an educational setting. However, the two races undoubtedly place different values on these fundamental skills even though they both function in our American society. Because of a difference in the interaction between the individual's educational environment and his cultural milieu it seems reasonable to assume that Indian children are not as cognizant of the ultimate need for these basic educational skills. Other than cultural differences, slow work habits, the nature of the test items, poor motivation, level of aspiration, and lack of interest may also account for some of the differences encountered. The question also arises as to whether the tests were equally fair to both groups. Undoubtedly, these conditions could have effected the results of this study.

V. Summary

The University of Kansas had conducted a fall achievement testing program for Indian and white children at the elementary, intermediate, and advanced grade levels for the United States Bureau of Indian Affairs during the period 1951-55. The California Achievement Test Batteries were used to obtain scores of educational achievement in the following basic skills: reading vocabulary, reading comprehension, arithmetic reasoning, arithmetic fundamentals, language, and spelling. When the mean scores of these tests were compared as to race, significant differences were found in favor of white children.

The sample was limited to all full-blooded Indian children and white children on whom a complete set of test scores was available. The distribution of the 14,888 subjects used in this study consisted of 5502 full-blooded Indian children and 9386 white children. The geographical locations of the administrative areas of the United States Bureau of Indian Affairs in which the testing was conducted were: 1) Phoenix, 2) Albuquerque, 3) Aberdeen, 4) Billings, 5) Muskogee, and 6) Anadarko.

Pearson product-moment correlations between the tests were computed for the various grade level groups by race and sex. The twelve intercorrelation matrices were factor analyzed by Thurstone's centroid method and the factors were subjected to oblique rotation by means of extended vectors in accordance with the concept of simple structure. Three factors resulted. With oblique rotation, the third factor was not significant for full-blooded Indian male children. Correlations between the reference vectors and the primary vectors were obtained. Factors A and C were identified as verbal factors and Factor B as a nu-

merical factor with highly significant correlations between the factors. This indicated a second-order factor common to all tests.

The purpose of this study was to compare the sets of factor patterns and factor structures by race, sex, and grade levels. Conclusions from the factor analyses follow:

1. Two factors were found for full-blooded Indian male children at each of the three grade level groups, while three factors were found at each grade level for full-blooded Indian female children and for male and female white children.
2. Higher factor loadings were found for full-blooded Indian male children than for any other race-sex group. The factor loadings ranged from .20 to .65 for full-blooded Indian children and from .21 to .57 for white children.
3. In most cases the factor loadings of a verbal nature were higher for full-blooded Indian male children.
4. In general, the factor loadings of a numerical nature were higher for full-blooded Indian children except on the arithmetic reasoning test where white female children at the advanced level had the highest loading.
5. The sets of factor patterns for all groups were more similar than they were different.
6. A common factor of the second-order domain was evident throughout all groups. This factor is probably Schooling.

FOOTNOTES

* This monograph is based on the author's dissertation completed at the University of Kansas. The data used in this study were obtained from the Indian Education Research Project conducted by the United States Bureau of Indian Affairs and the University of Kansas during the period 1951-55. Mr. L. M. Coombs, Education Specialist, United States Bureau of Indian Affairs, offered suggestions which were essential to the study. Dr. K. E. Anderson and Dr. A. H. Turney, Committee Members, read and evaluated the manuscript. Dr. R. N. Bradt, Department of Mathematics, provided helpful advice on the statistical and factorial analyses. Mr. C. Bradrick, Director of the Statistical Bureau, compiled the data for computation of the correlations. The author is indebted to these persons for their cooperation and assistance. The author is deeply indebted to his major advisor, Dr. E. Gordon Collister, not only for his excellent direction and advice, but also for his encouragement and inspiration.

** Now at the College of Education, University of Oklahoma.

1. The writer is indebted to the staff of the Statistical Bureau, University of Kansas, for carrying out the necessary machine operations for these factor analyses.
2. The following scale was used for the evaluation of factor loadings:

Insignificant loadings	.00 to .20
Low loadings	.20 to .30
Moderate loadings	.30 to .50
High loadings	.50 and above

BIBLIOGRAPHY

1. Anastasi, Anne. "Further Studies on the Memory Factor," Archives of Psychology, XXII (1932), pp. 1-60.
2. Anderson, K. E. The Relative Achievements of the Objectives of Secondary School Science in a Representative Sampling of Fifty-Six Minnesota Schools, unpublished Ph. D. dissertation, University of Minnesota, 1949.
3. Anderson, K. E., and others. The Educational Achievement of Indian Children. A Re-examination of the Question: How Well are Indian Children Educated? (Lawrence, Kansas: Haskell Institute, 1953).
4. Asch, S. E. "A Study of Change in Mental Organization," Archives of Psychology, XXVIII (1936), pp. 1-30.
5. Balinsky, B. "An Analysis of the Mental Factors of Various Age Groups from 9-60," Genetic Psychology Monographs, XXIII (1941) pp. 191-234.
6. Blackwell, A. M. "A Comprehensive Investigation into the Factors Involved in Mathematical Ability of Boys and Girls," British Journal of Educational Psychology, X (1940), pp. 143-53, 212-22.
7. Bryan, A. I. "Organization of Memory in Young Children," Archives of Psychology, XXIV (1934), pp. 1-56.
8. Carroll, J. B. "A Factor Analysis of Verbal Abilities," Psychometrika, VI (1941), pp. 279-307.
9. Clark, M. P. "Changes in Primary Mental Abilities with Age," Archives of Psychology, XLI (1944), pp. 1-30.
10. Clark, W. A Study of Factors Related to Mastery of Skills in Reading, Arithmetic, and Written Expression in Los Angeles County Senior High Schools, unpublished Ed. D. dissertation, University of Southern California, 1941.
11. Coleman, J. C. California Test of Mental Maturity: Summary of Investigations Number Three (Los Angeles: California Test Bureau, 1956), pp. 1-30.

12. Coombs, C. H. "A Factorial Study of Number Ability," Psychometrika, VI (1941), pp. 161-89.
13. Coombs, L. M., and others. The Indian Child Goes to School (Lawrence, Kansas: Haskell Institute, 1956).
14. Crysostom, F. "L'organization des Traits Mentaux avec l'Age," Annales de l'Assoc., Canadienne-pour l'Advancement des Sciences, IV (1938), pp. 1-124.
15. Cureton, E. E. "The Verbal Relations Factor and Vocabulary," American Psychologist, II (1947), pp. 286-87.
16. Cureton, E. E. "Vocabulary as an Arbitrary Definer of the Verbal Relations Factor," Invitational Conference on Testing Problems (New York: Educational Testing Service, 1948), pp. 47-9.
17. Curtis, H. A. "A Study of the Relative Effects of Age and Test Difficulty Upon Factor Patterns," Genetic Psychology Monograph, XL (1949), pp. 99-148.
18. Davis, F. B. "Fundamental Factors of Comprehension in Reading," Psychometrika, IX (1944) pp. 185-97.
19. Doppelt, J. E. "The Organization of Mental Abilities in the Age Range Thirteen to Seventeen," American Psychologist, V (1949), p. 242.
20. Fitzgerald, J. A., and Ludeman, W. W. "The Intelligence of Indian Children," Journal of Comparative Psychology, VI (1926), pp. 318-28.
21. French, J. W. "The Description of Aptitude and Achievement Tests in Terms of Rotated Factors," Psychometric Monograph (Chicago: University of Chicago Press, 1951).
22. Fruchter, B. Introduction to Factor Analysis (New York: Van Nostrand Publishing Co., 1954).
23. Garrett, H. E., and others. "The Age Factor in Mental Organization," Archives of Psychology, XXVI (1935), pp. 1-31.
24. Garrett, H. E. "Differentiable Mental Traits," Psychological Record, II (1938), pp. 259-98.
25. Garrett, H. E. "A Developmental Theory of Intelligence," American Psychologist, I (1946), p. 373.
26. Garth, T. R. "The Results of Some Tests on Full and Mixed Blood Indians," Journal of Applied Psychology, V (1921), pp. 359-72.
27. Garth, T. R. "A Comparison of Mental Abilities of Mixed and Full Blood Indians on a Basis of Education," Psychological Review XXIX (1922), pp. 221-36.
28. Garth, T. R. "A Comparison of the Intelligence of Mexican and Mixed and Full Blood Indian Children," Psychological Review, XXX (1923), pp. 388-401.
29. Garth, T. R. "The Intelligence of Full Blood Indians," Journal of Applied Psychology, IX (1925), pp. 382-89.
30. Garth, T. R. "The Intelligence of Mixed Blood Indians," Psychological Bulletin, XXIV (1927), pp. 179-80.
31. Guilford, J. P. Psychometric Methods (New York: McGraw Hill, 1954).
32. Hall, W. E., and Robinson, F. F. "An Analytical Approach to the Study of Reading Skills," Journal of Educational Psychology, XXXVI (1945), pp. 429-42.
33. Harris, C. W. "An Exploration of Language Skill Patterns," Journal of Educational Psychology, XXXIX (1948), pp. 321-36.
34. Helmer, Velma. The American Indian and Mental Tests, unpublished M. A. thesis, University of Kansas, 1925.
35. Holzinger, K. J., and Swineford, Frances. "A Study in Factor Analysis: The Stability of a Bi-factor Solution," Supplemented Educational Monograph, No. 48 (Chicago: University of Chicago Press, 1939), pp. 1-91.
36. Hunter, W. S. ans Sommermier, E. "The Relation of Degree of Indian Blood to Score on the Otis Intelligence Test," Journal of Comparative Psychology, II (1922), pp. 257-77.
37. Joseph, Alice, and others. The Desert People: A Study of the Papago Indians (Chicago: University of Chicago Press, 1949).
38. Kelley, T. L. Crossroads in the Mind of Man: A Study of Differentiable Mental Abilities (Los Angeles: Stanford University Press, 1928).
39. Klineberg, O. An Experimental Study of Speed and Other Factors in Racial Differences," Archives of Psychology, XV, No. 93 (1938), pp. 1-111.
40. Leighton, Dorothea C., and Kluckhohn, C. Children of the People (Cambridge: Harvard University Press, 1947).
41. MacGregor, G. Warriors Without Weapons: A Study of the Society and Personality Development of the Pine Ridge Sioux (Chicago: University of Chicago Press, 1945).
42. Meriam, L., and others. The Problem of Indian Administration (Baltimore: Johns Hopkins Press, 1928).
43. Peterson, S. How Well are Indian Children Educated? (Lawrence, Kansas: Haskell Institute, 1948).
44. Reichard, S. "Mental Organization and Age Level, Archives of Psychology, XLI (1944), pp. 1-29.
45. Richards, S. "Genetic Emergence of Factor Specificity," Psychometrika, VI (1941), pp. 37-42.
46. Rowe, E. C. "Five Hundred Sixty-Seven White and Two Hundred Sixty-Eight Indian Children Tested by the Binet-Simon Tests,"

Pedogogical Seminary, XXI (1914), pp. 454-68.

47. Schiller, B. "Verbal, Numerical, and Spatial Abilities of Young Children," Archives of Psychology, XXIV (1934), pp. 1-69.
48. Schneck, M. M. R. "The Measurement of Verbal and Numerical Abilities," Archives of Psychology, XVII (1929), pp. 1-69.
49. Swineford, Frances. "Growth in General and Verbal Bi-factors from Grades VII-X," Journal of Educational Psychology, XXXVIII (1947), pp. 257-72.
50. Swineford, Frances. "The Nature of the General, Verbal, and Spatial Bi-factors," Supplemental Education Monograph (Chicago: University of Chicago Press, 1948), pp. 1-72.
51. Swineford, Frances. "General Verbal and Spatial Bi-factors After Three Years," Journal of Educational Psychology, XL (1949), pp. 353-60.
52. Thompson, Laura and Joseph, Alice. The Hopi Way (Chicago: University of Chicago Press, 1944).
53. Thurstone, L. L. "Note on a Reanalysis of Davis' Reading Tests," Psychometrika, XI (1946), pp. 185-88.
54. Thurstone, L. L. Multiple-Factor Analysis (Chicago: University of Chicago Press 1947).
55. Wrigley, C., and Neuhaus, J. O. The Match of Two Sets of Factors, mimeographed paper presented to the American Psychological Association, San Francisco, September 5, 1955.

JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 28, Number 3, March 1960)

ATTITUDE OF KINDERGARTEN CHILDREN TOWARD SPECIALLY PREPARED STORY MATERIALS*

BILL SNIDER**
State University of Iowa

THE DEVELOPMENT of materials for helping the child become sensitive to the dynamics of behavior has been one of the aims of the Preventive Psychiatry Research Program in the Child Welfare Research Station at the State University of Iowa. Among the material prepared is a group of narratives designed to help the kindergarten child develop a causal orientation toward his social environment. If these stories are to be an effective medium for introducing the young child to a causal way of thinking, it seems necessary that the children should enjoy listening to them. The primary purpose of this study, therefore, was to discover how well kindergarten children liked to listen to and discuss causally oriented narratives specially written for their age. The statistical procedure chosen to determine this was the method of paired comparisons.

The purpose of the causally oriented story is to introduce the child to the differences between a non-causal or surface approach and a causal approach to human situations. In the narrative, a social situation is presented and someone begins to act in a non-causal way. The problem is then re-examined and the individual proceeds more causally after which the effects of this type of action are considered The reading of the story is accompanied by a discussion of differences between the two approaches and some of the probable effects of each.

Description of the Investigation

During the school year 1954-55, a number of teachers participated in the Preventive Psychiatry Program (5) by presenting causally oriented materials to their classes. The teachers had taken part in an intensified training program through summer workshops set up to train them to practice and teach the dynamic approach. The pupils in two of the kindergarten groups were chosen as subjects for this study. Class A consisted of 41 children in the morning and afternoon classes of teacher A, and Class B of 61 pupils in the morning and afternoon classes of teacher B.

In the method of paired comparisons, the subject is presented a number of different stimuli in pairs and is asked to judge which is greater with respect to some characteristic, such as, which of two objects is heavier, which is prettier, or which of two activities he prefers. In this particular investigation each child was asked which of two kindergarten activities he preferred. By statistical analysis of paired comparisons it is possible to assign a set of numbers to the different stimuli the subjects have been asked to judge.

Six different kindergarten activities were studied, namely: having music, working puzzles, having the teacher read ordinary or regular kindergarten stories, playing house, painting at the easel, and having the teacher read the causally oriented primary stories. These kindergarten activities were combined in all possible pairs. Each child was taken privately by a trained person[1] and asked which of each pair of activities he preferred.

The first few minutes with each child was spent in establishing rapport and seeing if the child understood the difference between the two types of stories being used as stimuli. The regular kindergarten stories, often read by the teacher, consisted of a pile of several books from the book table. The causally oriented stories were designated as the book without any pictures from which the teacher read to them; a copy of this book was on the table so that the child could recall it. A final check was made by asking the child if he knew the difference between the stories in this book without any pictures in it and the stories here—pointing to the books from the table. The administrator did not let the child see the books from the table other than the stack as the child was to respond to the books from the library table rather than to specific

* The preparation of this paper was supported by the Grant Foundation.
** Address: Preventive Psychiatry Research Program, Child Welfare Research Station.
1. Mrs. Helen Marie Clark, Elementary Supervisor for the Preventive Psychiatry Program, 1955-56, administered the test.

stories, and, likewise, the book containing the causally oriented stories was only shown to the child without permitting him to leaf through it.

The presentation of the pairs of stimuli were presented in such a manner as to adhere to the following principles as closely as possible (10, 11):

1. Maintain the greatest possible spacing between pairs involving any given member of the stimulus group.
2. Present a stimulus equally often first and second. This is referred to as taking account of "space errors."
3. Present stimuli as nearly in the same manner as is possible so as to avoid suggestion effects.

In order to insure the fullfillment of the third principle, the question for the presentation of each pair of objects was carefully planned and written in the instructions. Examples of these questions were:

1. Which would you rather do, listen to a story from this book (the causally oriented stories) or listen to a story from this book (regular stories)? (In each case only pointing to the book or books.)
2. Which would you rather do, have music time when you sing or work with a puzzle?
3. Which would you rather do, play house or paint at the easel?

Analyses of the Data

The first step in the analysis of preferences of the six kindergarten activities used in this study was the preparation of a proportion matrix showing the proportion of times each activity was preferred over its paired activity. Table I and Table II give the proportion matrices for the children in Classes A and B respectively. In these matrices each stimulus at the top is judged greater than each stimulus on the side by the corresponding proportion. For example, in Table I the proportion of the time "Music" was preferred over "Easel Painting" is .463. In Table II the proportion of time "Playing House" was preferred over "Regular Stories" is .623.

The analysis of the proportion matrices in Tables I and II is based upon the application of Thurstone's law of comparative judgment (13). This law is stated in the following form:

$$R_j - R_k = z_{jk}\sqrt{s_j^2 + s_k^2 - 2r_{jk}s_j s_k}$$

where R_j and R_k = mean psychological values characteristically attached to stimuli S_j and S_k respectively,
z_{jk} = standard measure distance or deviate from the mean of a unit normal distribution,
s_j and s_k = standard deviations of distributions R_{hj} and R_{hk} respectively, and
r_{jk} = coefficient of correlation between R_{hj} and R_{hk}.

The radical term is the standard deviation of the differences $R_{hj} - R_{hk}$ is expressed as a function of the terms under the radical.

In Thurstone's Case III, he assumes $r_{jk} = 0$, that there is no correlation between responses to any pair of stimuli. Under this assumption the law of comparative judgment reduces to the form

$$R_j - R_k = z_{jk}\sqrt{s_j^2 + s_k^2}$$

Scale-values in terms of preference for the six activities as judged by the children in Class A and Class B were analyzed under the assumption of Thurstone's Case III. This was done for Class A and Class B separately and for the combined group.

We find the z_{jk} values from a knowledge of the experimental proportions $p_{j>k}$. The z_{jk} is the z-value, given in a table of the normal curve, for the respective proportion, consider the proportion $p_{j>k}$ as the area under the normal curve. When the proportion p is greater than .500 it's z-value is positive; when p is less than .500 it's z-value is negative.

Table III gives the z_{jk} values for the respective proportions obtained for Class A as given in Table I. The sum column in Table III is used as a means of checking the computations. The sum of the sums should equal zero, and the sum of the column means should also equal zero.

Burros (1) proposed some convenient formulas for estimating the relative values for the standard deviations:

$$s_j = \frac{c}{\Sigma V_{z_j}}$$

where s_j = the standard deviation for the dispersion of stimulus S_j
V_{z_j} = the variance of deviations z_{jk} in each column of Matrix Z_5 from the column mean, $M_{z_{jk}}$, these values are found in Table III for Class A.
c is estimated by the formula

$$c \doteq \frac{n}{\Sigma \frac{1}{V_{z_j}}}$$

TABLE I

PROPORTION MATRIX FOR THE SIX KINDERGARTEN ACTIVITIES IN TERMS OF PREFERENCE AS JUDGED BY THE FORTY-ONE PUPILS IN CLASS A* (MATRIX P_A)

Kindergarten Activities	Regular Stories	Music	Playing House	Working Puzzles	Easel Painting	Causal Stories
Regular Stories	.500	.561	.561	.537	.610	.902
Music	.439	.500	.707	.561	.537	.732
Playing House	.439	.293	.500	.659	.561	.780
Working Puzzles	.463	.439	.341	.500	.659	.683
Easel Painting	.390	.463	.439	.341	.500	.659
Causal Stories	.098	.268	.220	.317	.341	.500
$\Sigma p_{j>k}$	2.329	2.524	2.768	2.915	3.208	4.256

*Kindergarten Activities have been arranged in order of increasing $\Sigma p_{j>k}$

TABLE II

PROPORTION MATRIX FOR THE SIX KINDERGARTEN ACTIVITIES IN TERMS OF PREFERENCE AS JUDGED BY THE SIXTY-ONE PUPILS IN CLASS B* (MATRIX P_B)

Kindergarten Activities	Regular Stories	Music	Playing House	Working Puzzles	Easel Painting	Causal Stories
Regular Stories	.500	.574	.623	.656	.787	.918
Music	.426	.500	.459	.459	.656	.705
Playing House	.377	.541	.500	.607	.525	.623
Working Puzzles	.344	.541	.393	.500	.627	.623
Easel Painting	.213	.344	.475	.328	.500	.443
Causal Stories	.082	.295	.377	.377	.557	.500
$\Sigma p_{j>k}$	1.942	2.795	2.827	2.927	3.697	3.812

*Kindergarten Activities have been arranged in order of increasing $\Sigma p_{j>k}$ across the columns.

TABLE III

Z-VALUES FOR THE RESPECTIVE PROPORTIONS OF THE SIX KINDERGARTEN ACTIVITIES IN TERMS OF PREFERENCE AS JUDGED BY THE FORTY-ONE PUPILS IN CLASS A (MATRIX Z_{5A})

Kindergarten Activities	Regular Stories	Music	Playing House	Working Puzzles	Easel Painting	Causal Stories	Sum
Regular Stories	.000	.154	.154	.093	.279	1.293	1.973
Music	- .154	.000	.545	.154	.093	.619	1.257
Playing House	- .154	- .545	.000	.410	.154	.772	.637
Working Puzzles	- .093	- .154	- .410	.000	.410	..476	.229
Easel Painting	- .279	- .093	- .154	- .410	.000	.410	- .526
Causal Stories	-1.293	- .619	- .772	- .476	- .410	.000	-3.570
Sum z_{jk}	-1.973	-1.257	- .637	- .229	.526	3.570	.000
Mean z_{jk}	- .329	- .210	- .106	- .038	.088	.595	.000

TABLE IV

ESTIMATION OF THE STANDARD DEVIATIONS OF THE RESPONSES TO THE SIX KINDERGARTEN ACTIVITIES AS JUDGED BY THE FORTY-ONE KINDERGARTEN CHILDREN IN CLASS A

Kindergarten Activities	Regular Stories	Music	Playing House	Working Puzzles	Easel Painting	Causal Stories	Sum
V_{z_j}	.1928	.0788	.1735	.0977	.0667	.1536	55.379
$1/V_{z_j}$	5.187	12.690	5.764	10.235	14.993	6.510	
s_j	.562	1.374	.624	1.108	1.624	.705	5.996
s^2_j	.316	1.888	.389	1.228	2.637	.497	

where n is the number of stimuli, and V_{z_j} is defined as above.

Substituting the z_{jk} and $M_{z_{jk}}$ values in Table III in the above formulas we can estimate the standard deviations for the six kindergarten activities as judged by the forty-one children in Class A. Table IV gives the statistics needed for computing Burro's formula for the Class A group.

The sum of the reciprocals of the variances is 55.379 and the number of stimuli is six, giving a c-value of .1083. As a check on the computations the s_j values should sum to n, the number of stimuli. This is due to the fact that this method assures a mean s_j of 1.000.

Table V contains the $(R_j - R_k)$ values which are computed according to the assumptions of Thurstone's Case III, $R_j - R_k = z_{jk}\sqrt{s_j^2 + s_k^2}$ and the R_j scale values are based upon these assumptions. The R_j value is a method of arbitrarily assigning the zero point to the lowest mean and adding the amount necessary to make the lowest mean .000 to each of the mean values. Either the mean values or the R_j values can be taken as scale values for the different kindergarten activities. Guilford (2) presents a method of locating a psychological meaningful zero point; however, this requires addtional data.

The sum column in Table V is used as a means of checking the computations in the table, the sum of the sums should equal zero, the sum of the columns means should also equal or approximate zero, the sum of the R_j's should equal n times the number added to each mean $M_{z_{jk}}$, or .353n, where n is the number of stimuli.

Given the scale values in Table V, the next problem is to determine how well the observed proportion of pupils' judgment agree with the expected proportion as determined by the above scale values. This test of internal consistency is made by utilizing Mosteller's method (8) by which a single chi-square test is made for the obtained and expected proportions. The first step is to find the expected z'_{jk} values. These are obtained by applying the following formula:

$$z'_{jk} = \frac{R_j - R_k}{\sqrt{s_j^2 + s_k^2}}$$

where the $R_j - R_k$ values in the numerator are obtained from the interpair scale differences given in the R_j row in Table V. The denominator is the square root of the sum of the paired s_j^2's in the last row of Table IV. These z'_{jk} values are given in Table VI. Only half the matrix is needed since the two halves are identical in absolute values. The diagonal cells have been ignored as there are no actual experimentally obtained values.

The expected proportions p', given in Table VII, are the areas under the normal curve for the respective z'_{jk} values in Table VI. If z' should be negative the area under the smaller portion of the curve is used, whereas when z' is positive the area under the larger portion is used.

The test for internal consistency as proposed by Mosteller (8) does not use the actual p and p' values since many of the proportions might be close to 1.00, which is the region where sampling distributions of proportions are not normal even for reasonably large samples. In order to compensate for this lack of normality, each proportion is transformed to a statistic θ, whose sampling distribution is normal. Theta being the angle whose sine is the $\sqrt{p}$, or $\theta = \arcsin\sqrt{p}$. Tables VIII and IX give the θ-values for the obtained and experimental proportions respectively. The chi-square form applicable for this test of internal consistency is

$$\text{chi-square} = \frac{N}{821}\ \Sigma(\theta - \theta')^2$$

where N is the number of judgments per stimulus pair, θ is the $\arcsin\sqrt{p}$ and θ' is the $\arcsin\sqrt{p'}$. Mosteller (8) gives the degree of freedom for the above chi-square as

$$df = \frac{(n - 1)(n - 4)}{2}$$

where n is the number of stimuli. The sum of the squared differences between the observed and expected theta is 259.6927, and the number of pupils is forty-one; therefore

$$\text{Chi-square} = \frac{41(259.6927)}{821} \qquad df = 5$$

$$\text{Chi-square} = 12.969$$

Since the chi-square is not significant at the one percent level one could assume internal consistency between the observed and expected proportions for the preference of the six kindergarten activities as judged by the forty-one children in Class A. The scale-values for these activities are the R_j values given in Table V, namely: Regular Stories .000, Music .017, Playing House .259, Working Puzzles .266, Easel Painting .524, and Causal Stories .1.053. The zero point is arbitrarily assigned.

Analysis of the preference of the six kindergarten activities for Class B, using Thurstone's Case III assumptions yielded slightly different results as obtained for Class A. Table II gives the proportion of time each activity was preferred over its respective pair. The transformation of these proportions into $R_j - R_k$ values where $R_j - R_k = z_{jk}\sqrt{s_j^2 + s_k^2}$ and the resultant scale-values are presented in Table X. Testing of the obtained

TABLE V

SCALE SEPARATIONS BETWEEN PAIRS OF SIX KINDERGARTEN ACTIVITIES AS JUDGED IN TERMS OF PREFERENCE BY THE FORTY-ONE KINDERGARTEN PUPILS IN CLASS A, ASSUMING CASE III (MATRIX Z_{3A})

Kindergarten Activities	Regular Stories	Music	Playing House	Working Puzzles	Easel Painting	Causal Stories	Sum
Regular Stories	.000	.229	.129	.116	.479	1.166	2.119
Music	- .229	.000	.822	.272	.198	.956	2.019
Playing House	- .129	- .822	.000	.522	.268	.726	.565
Working Puzzles	- .116	- .272	- .252	.000	.806	.625	.521
Easel Painting	- .479	- .198	- .268	- .806	.000	.726	-1.025
Causal Stories	-1.166	- .956	- .726	- .625	- .726	.000	-4.199
Sum z_{jk}	-2.119	-2.019	- .565	- .521	1.025	4.199	.000
Mean z_{jk}	- .353	- .336	- .094	- .087	.171	.700	.001
R_j	.000*	.017	.259	.266	.524	1.053	2.118

*Arbitrarily assigned zero point

TABLE VI

EXPECTED SCALE SEPARATIONS z' BETWEEN THE SIX KINDERGARTEN ACTIVITIES DERIVED FROM SCALE VALUES OBTAINED BY THE ASSUMPTIONS OF CASE III (MATRIX Z'_A)

Kindergarten Activities	Music	Playing House	Working Puzzles	Easel Painting	Causal Stories
Regular Activities	.011	.308	.214	.305	1.167
Music		.160	.141	.238	.671
Playing House			.006	.152	.844
Working Puzzles				.131	.599
Easel Painting					.299

TABLE VII

EXPECTED PROPORTIONS p' DERIVED THROUGH z' FROM SCALE VALUES FOR SIX KINDERGARTEN ACTIVITIES, AS JUDGED BY PUPILS IN CLASS A, UNDER THE ASSUMPTION OF CASE III (MATRIX P'_A)

Kindergarten Activities	Music	Playing House	Working Puzzles	Easel Painting	Causal Stories
Regular Stories	.504	.621	.585	.620	.878
Music		.564	.536	.594	.749
Playing House			.502	.560	.801
Working Puzzles				.552	.725
Easel Painting					.618

TABLE VIII

OBTAINED PROPORTIONS FOR THE SIX KINDERGARTEN ACTIVITIES CONVERTED INTO ANGLES IN TERMS OF DEGREES (θ) BY MEANS OF THE FUNCTION $\theta = \text{ARCSIN}\sqrt{p}$ (MATRIX θ)

Kindergarten Activities	Music	Playing House	Working Puzzles	Easel Painting	Causal Stories
Regular Stories	48.50	48.50	47.12	51.35	71.76
Music		57.23	48.50	47.12	58.82
Playing House			54.27	48.50	62.03
Working Puzzles				54.27	55.73
Easel Painting					54.27

θ values for Class A.

TABLE IX

EXPECTED PROPORTIONS FOR THE SIX KINDERGARTEN ACTIVITIES CONVERTED INTO ANGLES IN TERMS OF DEGREES (θ') BY MEANS OF THE FUNCTION $\theta' = \text{ARCSIN}\sqrt{p'}$ (MATRIX θ')

Kindergarten Activities	Music	Playing House	Working Puzzles	Easel Painting	Causal Stories
Regular Stories	45.23	52.00	49.89	51.94	69.56
Music		48.68	48.22	50.42	59.93
Playing House			45.11	48.45	63.51
Working Puzzles				47.98	58.37
Easel Painting					51.83

θ' values for Class A.

TABLE X

SCALE SEPARATIONS BETWEEN PAIRS OF SIX KINDERGARTEN ACTIVITIES AS JUDGED IN TERMS OF PREFERENCE BY THE SIXTY-ONE KINDERGARTEN PUPILS IN CLASS B, ASSUMING CASE III (MATRIX Z_{3B})

Kindergarten Activities	Regular Stories	Music	Playing House	Working Puzzles	Easel Painting	Causal Stories	Sum
Regular Stories	.000	.222	.648	.416	.966	.740	2.993
Music	- .222	.000	- .239	- .152	.644	.632	.663
Playing House	- .648	.239	.000	.610	.147	.645	.993
Working Puzzles	- .416	.152	- .610	.000	.665	.318	.109
Easel Painting	- .966	- .644	- .147	- .665	.000	- .171	-2.593
Causal Stories	- .741	- .632	- .645	- .318	.171	.000	-2.165
Sum z_{jk}	-2.993	- .663	- .993	- .109	2.593	2.165	.000
Mean z_{jk}	- .499	- .110	- .166	- .018	.432	.361	.000
R_j	.000*	.389	.333	.481	.931	.860	2.994

*Arbitrarily assigned zero point

TABLE XI

PROPORTION MATRIX FOR THE SIX KINDERGARTEN ACTIVITIES IN TERMS OF PREFERENCE AS JUDGED BY THE COMBINED CLASSES (MATRIX P_C)

Kindergarten Activities	Regular Stories	Music	Playing House	Working Puzzles	Easel Painting	Causal Stories
Regular Stories	.500	.569	.598	.608	.716	.912
Music	.431	.500	.559	.500	.608	.716
Playing House	.402	.441	.500	.627	.539	.686
Working Puzzles	.392	.500	.373	.500	.667	.647
Easel Painting	.284	.392	.461	.333	.500	.529
Causal Stories	.088	.284	.314	.353	.471	.500
$p_{\hat{j}\hat{k}}$	2.097	2.686	2.805	2.921	3.501	3.990

TABLE XII

SCALE SEPARATIONS BETWEEN PAIRS OF SIX KINDERGARTEN ACTIVITIES AS JUDGED IN TERMS OF PREFERENCE BY THE 102 PUPILS OF THE COMBINED CLASSES, ASSUMING CASE III (MATRIX Z_{3C})

Kindergarten Activities	Regular Stories	Music	Playing House	Working Puzzles	Easel Painting	Causal Stories	Sum
Regular Stories	.000	.259	.327	.289	.900	.702	2.567
Music	- .259	.000	.281	.000	.571	.849	1.442
Playing House	- .327	- .281	.000	.513	.193	.639	.737
Working Puzzles	- .289	.000	- .513	.000	.779	.398	.375
Easel Painting	- .900	- .571	- .193	- .779	.000	.115	2.328
Causal Stories	- .792	- .849	- .639	- .398	- .115	.000	-2.793
Total	-2.567	-1.442	- .737	- .375	2.328	2.793	.000
Mean z_j	- .428	- .240	- .123	- .062	.388	.466	.001
R_j	.000*	.188	.305	.366	.816	.894	2.569

*Arbitrarily assigned zero point

TABLE XIII

SCALE VALUES FOR THE SIX KINDERGARTEN ACTIVITIES COMPUTED BY THE METHOD OF PAIRED COMPARISONS DERIVED BY THE ASSUMPTIONS OF THURSTONE'S CASE III

Kindergarten Activities	Class A	Class B	Combined Class A and B
Causal Stories	1.053	.860	.894
Easel Paintings	.524	.931	.816
Working Puzzles	.266	.481	.366
Playing House	.259	.333	.305
Music	.017	.389	.188
Regular Stories	.000*	.000*	.000*

*Arbitrarily assigned zero point

and expected proportions by Mosteller's method showed internal consistency of the data at the one percent level, the level chosen to reject the hypothesis. The scale-values computed for the six kindergarten activities as derived by means of paired comparisons and judged by the sixty-one pupils in Class B were Regular Stories .000, Playing House .333, Music .389, Working Puzzles .481, Causal Stories .860, and Easel Painting .931. Two pairs of activities reversed position when compared with scale-values for the kindergarten activities as judged by the children in Class A, namely, Music and Playing House, and Easel Painting and the Causal Stories.

As another part of the analysis of the investigation, the data for Class A and Class B were pooled and scale-values for the combined group were calculated. Table XI gives the obtained proportions for the combined group.

Thurstone's Case III assumptions were applied to the proportions obtained by pooling the data. The $R_j - R_k$ values under these assumptions are presented in Table XII plus the scale-values for each of the kindergarten activities.

The chi-square test for internal consistency was not significant at the one percent level. The scale-values thus obtained by Thurstone's Case III assumptions for the combined classes were Regular Stories .000, Music .188, Playing House .305, Working Puzzles .366, Easel Painting .816, and Causal Stories .894. The zero point is an arbitrarily assigned value.

Conclusions and Summary

The purpose of this experiment was to find out how well kindergarten children liked to listen to the causally oriented stories specially written for their age. Subjects were forty-one kindergarten children in Class A and sixty-one in Class B. Thurstone's Case III assumptions were used. After comparing this activity with five other kindergarten activities by means of paired comparisons, the results showed that listening to the prepared causal stories was preferred by Class A and the combined group over the other five kindergarten activities, namely, having Music, Working Puzzles, Painting at the Easel, Playing House, or Having Regular Kindergarten Stories read to them. The significance of internal consistency was tested by using the one percent level of rejection. A similar analysis of the paired comparison data for Class B showed a reversal in the preference of Causal Stories and Easel Painting. Easel Painting rated first and causal narratives, second.

Thus, using the materials designed to help the child develop a causal orientation toward his social environment ranked well with other activities included in this investigation.

Table XIII is a summary of the scale values assigned to the six kindergarten activities as judged by pupils in Class A, Class B, and the Combined Classes A and B, using Thurstone's Case III assumptions for scaling the paired comparison data.

REFERENCES

1. Burros, R. H. "The Application of the Method of Paired Comparisons to the Study of Reaction Potentials," Psychological Review LVIII (1951), pp. 60-66.
2. Guilford, J. P. "The Method of Paired Comparisons as a Psychometric Method," Psychological Review, XXXV (1928), pp. 494-506.
3. Guilford, J. P. "Some Empirical Tests of the Method of Paired Comparisons," Journal of General Psychology, V (1931), pp. 64-76.
4. Guilford, J. P. Psychometric Methods, Second Edition (New York: McGraw Hill Book Co., 1954), pp. 154-77.
5. Levitt, E. E., and Ojemann, R. H. "The Aims of Preventive Psychiatry and 'Causality' as a Personality Pattern," Journal of Psychology, XXXVI (1953), pp. 393-400.
6. Mosteller, F. "Remarks on the Method of Paired Comparisons: I. The Least Squares Solution Assuming Equal Standard Deviations and Equal Correlations," Psychometrika, XVI (1951), pp. 3-9.
7. Mosteller, F. "Remarks on the Method of Paired Comparisons: II. The Effect of an Aberrant Standard Deviation When Equal Standard Deviations and Equal Correlations are Assumed," Psychometrika, XVI (1951), pp. 203-6.
8. Mosteller, F. "Remarks on the Method of Paired Comparisons: III. A Test of Significance for Paired Comparisons When Equal Standard Deviations and Equal Correlations are Assumed," Psychometrika, XVI (1951), pp. 207-218.
9. Ojemann, R. H., and others. "The Effects of a 'Causal' Teacher-Training Program and Certain Curricular Changes on Grade School Children," Journal of Experimental Psychology, XXXVI (1953), pp. 393-400.
10. Ross, R. T. "Optimum Orders for the Presentation of Pairs in the Method of Paired Comparisons," Journal of Educational Psychology, XXV (1934), pp. 375-82.
11. Ross, R. T. "Optimal Orders in the Method of Paired Comparisons," Journal of Experimental Psychology, XXV (1939), pp. 414-24.
12. Thurstone, L. L. "Psychophysical Analysis," American Journal of Psychology, XXXVIII (1927), pp. 368-89.
13. Thurstone, L. L. "A Law of Comparative

Judgment," Psychological Review, XXXIV (1927), pp. 273-86.

14. Thurstone, L. L. "The Method of Paired Comparisons for Social Values," Journal of Abnormal and Social Psychology, XXI (1927), pp. 384-400.

JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 28, Number 3, March 1960)

ANALYSIS OF MMPI DATA FOR CLASSIFICATION PURPOSES BY MULTIVARIATE STATISTICAL TECHNIQUES*

PETER P. REMPEL
State University of Iowa

THE PURPOSE of this paper is to present the steps involved in the application of multivariate analysis to the assignment—with a minimum error of classification—of N people to q criterion groups on the basis of all information provided by k intercorrelated measures.

The study itself stems from the writer's attempt to identify as potential delinquents or nondelinquents (q = 2), a population of ninth-grade high-school boys (N = 1802) to whom the MMPI had been administered by Hathaway (3) in 1947-1948, and for whom police and/or court records were available through 1952. Severity of misdemeanor made possible a classification of delinquency into eight levels. Because of inter-group similarity of MMPI profiles between nondelinquents and delinquents at levels one and two, the latter (N = 363) have been omitted from this study (see Table I).

Variables used in one part of the study, referred to as the main study, consisted of the scores on the F, Hs, Sc, Mf, and Pd scales of the MMPI (k = 5); in another part, the school study, they consisted of the Ma, Sc, Mf, and Pd scales of the MMPI; honor point ratio, and numbers of days absent in Grade VIII (k = 6).

In the main study the MMPI profiles of 351 boys who had achieved delinquency status by 1952 were divided by random procedures into two groups: Sample A (N = 175) which was used in the development of the classification formula, and Sample B (N = 176), set aside for cross-validation purposes. Two samples of 175 profiles each were similarly selected from the nondelinquent set of MMPI profiles. In the school study both delinquent and nondelinquent groups were divided into Samples A' and B' and four sub-groups (N = 91) were thus similarly obtained.

Tests of Assumption

Numerical values obtained for the statistics β_1 and β_2—the former used as a measure of skewness, the latter as a measure of kurtosis—indicated that roughly the distributions of only half of the variables tested in either study (in Samples A and A') satisfied the assumption of normality. Somewhat more satisfying results were obtained from tests of homogeneity of variances and covariances. Thus approximately 30 percent of the tests indicated lack of homogeneity of variances at a statistically significant level, and 20 percent indicated lack of homogeneity of covariances. The effect of nonnormality upon the distribution of D^2 and the LDF statistic is not known. A test of the assumption of linearity of regression (4: 240-245) was omitted because the number of variables involved made the amount of labor required too excessive. Inspection of correlation charts of selected pairs of variables made it appear unlikely that anything other than linearity existed. Because cross-validation has been included, the application of multivariate analysis was thought to be justified even though not all of the assumptions have been fully met. But in drawing conclusions from this study, one must recognize the fact that not all of the assumptions for the application of multivariate procedures were fulfilled.

Discussion of Statistical Techniques

The generalized distance function contributes both to the problem of testing significance and of estimation. The five (or six) variables used in the main and school studies respectively, may be considered as defining a space of five (or six) dimensions so that any individual with assigned scores can be represented by a point in such a space. The individuals belonging to any group will then be represented by a cluster of points round another point which represents the mean values of these scores. If the mean values of the two groups are close together, then the clusters of points corresponding to them will overlap to some extent. Rao (6) has shown that the degree of overlap between any two

*Submitted to the Graduate School of the University of Minnesota in partial fulfillment of the requirements for the Ph. D. degree.

TABLE I

DISTRIBUTION OF VALID MMPI PROFILES BY DELINQUENTS, LEVEL OF DELINQUENCY, AND NON-DELINQUENTS

Classification	Number	Percent (Based on N = 1802)	Percent (Based on N = 714)
All Delinquents	714	39.62	
Levels 1 and 2	363	20.15	50.84
Level 3	231	12.82	32.35
Level 4	47	2.61	6.58
Level 5	31	1.72	4.34
Level 6	22	1.22	3.08
Level 7	12	0.67	1.68
Level 8	8	0.44	1.12
Total Level 3-8	351	19.47	49.16
Non-delinquents	1088	60.37	.....
Total (All delinquents and non-delinquents)	1802	100.00	.....

TABLE II

A SUMMARY COMPARISON OF CLASSIFICATION EFFICIENCY FOR THE LDF PROCEDURE WITH CHANCE EXPECTATION FOR SAMPLES A AND B IN THE MAIN STUDY BY DELINQUENT AND NON-DELINQUENT GROUPS WITHIN THE SAMPLE

Samples Used		Correct Classification by Formula N	Correct Classification by Formula %	Correct Classification by Chance N	Correct Classification by Chance %	Percentage Improvement Over Chance
Sample A - Delinquents	N = 175	120	68.6	87.5	50.0	18.6**
Sample A - Non-delinquents	N = 175	115	65.7	87.5	50.0	15.7**
Total Sample A	N = 350	235	67.1	175.0	50.0	17.1**
Sample B - Delinquents	N = 176	124	70.5	88.0	50.0	20.5**
Sample B - Non-delinquents	N = 175	103	58.9	87.5	50.0	8.9*
Total Sample B	N = 351	227	64.7	175.5	50.0	14.7**
All Groups Total	N = 701	462	65.9	350.5	50.0	15.9**

* Significant at the 5 percent level
**Significant at the 1 percent level

such clusters can be measured by the quantity

$$D^2 = \sum_{i=1}^{i=5(6)} \sum_{j=1}^{j=5(6)} a^{ij} d_i d_j$$

where $d_1 d_2; \ldots, d_5$ (or d_6) are the differences between the mean values of the measurements in the two groups, and a^{ij} are the elements of the matrix reciprocal to the within group dispersion matrix (a_{ij}). The higher the value of D^2, the smaller the amount of overlap, so that this measure may be regarded as the "distance" between the groups.

The calculation of D^2 is given in the basic form as

$$D^2 = \Sigma_i \Sigma_j S^{ij} d_i d_j \qquad (1)$$

S^{ij} = elements of the matrix inverse to S_{ij}, the dispersion or within groups variance-covariance matrix.

$$d_i = \overline{X}_i - \overline{X}_i'$$

$$d_j = \overline{X}_j - \overline{X}_j'$$

where $\overline{X}_i$ represents the sample mean of the ith variable and $\overline{X}_i'$ is the sample mean of a second category of the ith variable.

The "best" (unbiased) estimate of the sample variance used in the matrix for the ith test is

$$\sigma_i^2 = \frac{1}{n}\left[\sum_{k=1}^{N_x} X_{ik}^2 - N_x \overline{X}_i^2 + \sum_{k=1}^{N_y} Y_{ik}^2 - N\overline{Y}_i^2\right] \qquad (2)$$

and the unbiased estimate of the covariance becomes

$$\rho_{ij} = \rho_{ji} = \frac{1}{n}\left[\sum_{k=1}^{N_x} X_{ik} X_{jk} - N_x \overline{X}_i \overline{X}_j + \sum_{k=1}^{N_y} Y_{ik} Y_{jk} - N_y \overline{X}_i \overline{X}_j\right] \qquad (3)$$

where

X_i $(i = 1, 2, \ldots, r)$ = the random variable defined by the ith test (or scale) in the first group,

Y_i $(i = 1, 2, \ldots, r)$ = the random variable defined by the ith test (or scale) in the second group,

X_{ik} = the score of the kth individual in the first group on the ith test.

Y_{ik} = the score of the kth individual in the second group on the ith test,

$\overline{X}_i$ = the sample mean of the ith test in the first group, and

$\overline{Y}_i$ = the sample mean of the ith test in the second group.

Rao (9) met the problem of eliminating the correlation between variables by suggesting a method of transforming the original set of data to an orthogonal set of variates. The most efficient method of accomplishing this purpose is the application of pivotal condensation to the calculation of D^2, and is the method used in this study. A systematic and straight-forward illustration of the use of Rao's pivotal condensation technique is found in a paper by Collier (2).

The application of pivotal condensation to the problem of transforming correlated variates to a new orthogonal set reduces the covariance elements in the S_{ij} variance-covariance matrix to zero, thus eliminating the element of correlation from matrix calculations. The general distance function may then be expressed as

$$D^2 = \frac{\Sigma^* d_i^2}{^*S_{ii}}, \qquad (4)$$

in which $^*S_{ii}$, the variance of the transformed ith variable is found easily in the process of pivotal condensation, and each $\frac{^*d_i^2}{^*S_{ii}}$ represents the contribution of the ith variable to the total D^2.

Significance of Added Variables

Rao (7, 8, 9) has considered the following approximate test of significance which can be used for testing both the significance of the total D^2 value as well as of the addition to D^2 of each variable separately:

$$T = \frac{(f-p-q+1)}{q}\left[\frac{n_1 n_2 (D^2_{p+q} - D^2_p)}{f(n_1+n_2)+n_1 n_2 D^2_p}\right] \qquad (5)$$

where f = the degrees of freedom of the dispersion matrix of variance and covariances between the ith and jth characters

p = number of variables in D_p^2

q = number of variables added to form D_{p+q}^2

n_1 = size of sample or groups from the first population considered

n_2 = size of sample from the second population

T has the F-distribution with q and f-p-q+1 degrees of freedom. Use of this value enables the research worker to determine the effect of adding variables to the matrix.

The Linear Discriminant Function (LDF)

Let X_{ijk} be a typical observation where i = 1, 2 stands for the sample, j for the trait measured, and k for the individual. Then the value Y, would be determined by the function

$$Y_{ik} = \lambda_1 X_1 + \lambda_2 X_2 + \ldots + \lambda_p X_p \tag{6}$$

and

$$\overline{Y}_{ik} = \frac{i}{n_i} \Sigma_i Y_{ik} \tag{7}$$

where Σ_i denotes summation over the ith sample (i = 1, 2). The problem is to maximize the ratio

$$\frac{\overline{Y}_2 - \overline{Y}_1}{\sqrt{\Sigma_1 (Y_{1k} - \overline{Y}_1)^2 + \Sigma_2 (Y_{2k} - \overline{Y}_2)^2}} \tag{8}$$

which may be written as

$$\frac{\left(\sum_{j=1}^{p} \lambda_j d_j\right)^2}{\sum_{j=1}^{p} \sum_{m=1}^{p} \lambda_j \lambda_m S_{jm}} \tag{9}$$

where $d_j = \frac{1}{n_2} \cdot \Sigma_2 X_{2jk} - \frac{1}{n_1} \Sigma_1 X_{1jk}$ for the jth quantity (j = 1, 2, . . . , p): and S_{jm} is the pooled sum of products of deviations from the sample traits j and m. Maximization of the above ratio requires essentially setting the derivative of the ratio equal to zero, and we obtain

$$\frac{\sum_{m=1}^{p} \sum_{j=1}^{p} \lambda_j \lambda_m S_{jm}}{\sum_{j=1}^{p} \lambda_j d_j} \cdot dr = \sum_{j=1}^{p} \lambda_j \lambda_{jr}, \ (r = 1, 2, \ldots, p) \tag{10}$$

Since the first factor in the above equation will be the same for all equations it can be ignored. Therefore, the coefficients required are obtained through the solutions of the normal equations

$$\left.\begin{aligned} S_{11}\lambda_1 + S_{12}\lambda_2 \cdots + S_{1p}\lambda_p &= d_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ S_{p1}\lambda_1 + S_{p2}\lambda_2 \cdots + S_{pp}\lambda_p &= d_p \end{aligned}\right] \tag{11}$$

The diagonal cells in the above matrix represent variances, while all other cells consist of the appropriate covariances. D^2 is represented by the expression

$$D^2 = \sum_{i=1}^{p} \lambda_i d_i \tag{12}$$

and the discriminant function by

$$LDF = \sum_{i=1}^{p} \lambda_i X_i \tag{13}$$

Classification Procedures

The application of Fisher's discriminant function to the problem of classification into one or two known multivariate normal populations is discussed by Anderson (1). Suppose that an individual I with a set of measurements $X_1, \ldots, X_q$ has been randomly drawn from one of two populations Π_1 and Π_2, with probability distributions of the measurements $f_1(X_1, \ldots, X_q)$, respectively. Suppose further that $X_1, \ldots, X_q$ has a joint normal distribution with means in Π_1 of $EX_i = u_i$ (1), and in Π_2 of $EX_i = u_i$ (2). Let the common set of variances and covariances be

$$\sigma_1^2, \ldots, \sigma_{pi}^2 \ \rho_{12}, \rho_{13}, \ldots \rho_{p-1, p}.$$

It can then be shown that the best regions of classification are of the general form

$$R_1 : \sum_{i=1}^{p} \lambda_i X_i > \sum_{i=1}^{p} \lambda_i \tfrac{1}{2} (u_i^{(1)} + u_i^{(2)}) + \log k \tag{14}$$

$$R_2 : \sum_{i=1}^{p} \lambda_i X_i < \sum_{i=1}^{p} \lambda_i \tfrac{1}{2} (u_i^{(1)} + u_i^{(2)}) + \log k \tag{15}$$

where $\lambda_1, \ldots, \lambda_p$ is the solution of

$$\sum_{j=1}^{p} \sigma_i \sigma_j \rho_{ij} \lambda_j = u_i^{(1)} - u_i^{(2)} \tag{16}$$

and $\log k = \log \frac{L(1 \mid 2) p_2}{L(2 \mid 1) p_1}$

where L(2 | 1) is defined as the loss incurred in classifying individual I into Π_2 when in fact he belongs to Π_1, and L(1 | 2) as the loss incurred in classifying individual I into Π_1 when in fact he belong to Π_2.

The first term of Formulas 14 and 15 is the same as that to which the LDF simplifies (see Formula 13) if the population expectations and covariance matrix are known. Therefore if k = 1, i.e., if the probability of belonging to one group is equal to coming from the other, then log k = 0, and the procedure is reduced to a comparison of the LDF of individual observations with the LDF of the averages of the respective means. If a priori probabilities are not specified, log k is found

so that the risk when I is from Π_1 is equal to the risk when I is from Π_2. For this purpose let

$$\sum_{i}^{q} \lambda_i X_i - \tfrac{1}{2} \sum_{i}^{q} \lambda_i (u_i^{(1)} + u_i^{(2)}) = U(X_i, \ldots, X_q) \quad (17)$$

Since a linear function of normally distributed variables is again normally distributed, the distribution of U is likewise normal, with the mean and variances found as follows:

$$E(U) = \sum_{i}^{q} \lambda_i \, E(X_i) - \tfrac{1}{2} \sum_{i}^{q} \lambda_i (u_i^{(1)} + u_i^{(2)}) \quad (18)$$

and $$\sigma_U^2 = \sum_{i}^{q} \lambda_i (u_i^{(1)} - u_i^{(2)}) \quad (19)$$

Let $$\sigma_U^2 = a \quad (20)$$

Then if I is from Π_1, $E(X_i) = u_i^{(1)}$ and

$$E(U) = \tfrac{1}{2} \sum_{i}^{q} \lambda_i (u_i^{(1)} - u_i^{(2)}) = \tfrac{1}{2}a \quad (21)$$

where a is the "distance" between Π_1 and Π_2 or the variance of U. If I is from Π_2, $E(X_i) = u_i^{(2)}$, and

$$E(U) = \sum_{i}^{q} \lambda_i (u_i^{(1)} - u_i^{(2)}) = -\tfrac{1}{2}a \quad (22)$$

The probability of misclassifying an individual belonging to Π_1 is then determined by referring the following relation to a normal distribution table:

$$P\left(Z < \frac{\log_e k - \frac{1}{2}a}{\sqrt{a}} \right) \quad (23)$$

The probability of misclassifying an individual belonging to Π_2 is determined by referring the following relation to a normal distribution table:

$$P\left(Z > \frac{\log_e k - \frac{1}{2}a}{\sqrt{a}} \right) \quad (24)$$

In both cases $k = \Pi_2 \; \Pi_1$. If $\Pi_2 = \Pi_1$ then $k = 1$, and $\log k = 0$. Then the formula for calculating the theoretical probability of misclassification is reduced to

$$P\left(Z > \frac{o - (-\frac{1}{2}a)}{\sqrt{a}} \right) \quad (25)$$

The overall probability of misclassification is then the sum of products of the population proportions in each group and the respective misclassification probabilities.

Test of Effectiveness of Linear Discriminant Functions

Lubin (5) has given us the tests of the effectiveness of the LDF in classifying individuals correctly. In a two-contingency table let the columns A_1 and A_2 represent the observed categories while the rows B_1 and B_2 represent the predicted categories. If perfect discrimination is achieved, only the diagonal cells will be filled. For pure chance discrimination

$$n_i = \frac{A_i B_j}{N} \quad (26)$$

where

- n_{ij} = frequency of the ith row and the jth column
- A_i = marginal total of the ith row
- B_j = marginal total of the jth row
- N = total number of individuals involved.

In case the value of chi-square computed on the diagonal values is found to be significant (and a chi-square test is a necessary but not sufficient test to demonstrate classification beyond chance), then the one-tailed t-test for large values of N, may be used to demonstrate the effectiveness of the LDF:

$$t = \sqrt{\frac{(o - e)\sqrt{N}}{e(N - e)}} \quad (27)$$

where o denotes the number of individuals classified correctly, and e the number of correct classifications expected by chance.

Computational Procedures

In the main study the variance-covariance matrix representing the equation

$$\sum_{j=1}^{5} \sigma_i \sigma_j \rho_{ij} \lambda_j = u_i^{(1)} - u_i^{(2)}, \; i = 1, \ldots, 5$$

was found to be as shown at top of next page.

This set of simultaneous linear equations for Sample A was solved by the method of pivotal condensation. This method reduces the original set of variables to a new orthogonal set of uncorrelated variables, and makes possible a test of significance of the contribution of each variable to D^2. The numerical values of the beta weights and the individual contribution of variables to D^2 are as shown in the middle of the next page.

The value of D^2 computed from the product of these weights and the difference in means for each variable was found to be .9173 (i.e., the same that

	F	Hs	Sc	Mf	Pd	$u_i^{(1)} - u_i^{(2)}$
F	11.9238 λ_1 +	1.9755 λ_2 +	17.1353 λ_3 +	3.5074 λ_4 +	8.4253 λ_5 =	1.5029
Hs	1.9755 λ_1 +	77.1411 λ_2 +	38.6356 λ_3 +	19.5887 λ_4 +	30.8370 λ_5 =	2.2743
Sc	17.1353 λ_1 +	38.6256 λ_2 +	104.1166 λ_3 +	28.5664 λ_4 +	50.1619 λ_5 =	4.9714
Mf	3.5074 λ_1 +	19.5887 λ_2 +	28.5664 λ_3 +	72.7225 λ_4 +	16.9858 λ_5 =	-3.5029
Pd	8.4253 λ_1 +	30.8370 λ_2 +	50.1619 λ_3 +	16.9858 λ_4 +	99.6932 λ_5 =	6.6629

Scale	Beta Weights	Contribution of Variables to D^2
F	λ_1 = .0695	.1894**
Hs	λ_2 = .0116	.0534*
Sc	λ_3 = .0256	.0556*
Mf	λ_4 = -.0782	.3750**
Pd	λ_5 = .0578	.2439**
		D^2 = .9173

was obtained by adding the contributions of variables to D^2 above).

The first part of the expression

$$U = \sum_{i=1}^{p} \lambda_i X_i - \sum_{i=1}^{p} \lambda_i \tfrac{1}{2}(u_i^{(1)} + u_i^{(2)})$$

is obviously the LDF, and for any set of scores in Sample A, its numerical value will be written in terms of the beta weights for D^2 as follows:

$$LDF = \sum_{i=1}^{p} \lambda_i X_i = .0695X_1 + .0116X_2 + .0256X_3 - .0782X_4 + .0578X_5$$

The second part of the expression represents the middle point between the sums of the weighted means of the two criterion groups and is therefore a constant quantity, c. For Sample A the value of c is the sum of the products of the beta weights and the sum of means, or

$$c = \tfrac{1}{2}[(.0695)(12.309) + (.0116)(103.4514) + (.0256)(119.0628) + (-.0782)(103.6629) + (.0578)(121.6914)] = 2.0169$$

Also, since

$$\sigma_U^2 = \sum_{i}^{q} \lambda_i (u_i^{(1)} - u_i^{(2)}) = a,$$

and

$$\sum_{i}^{q} \lambda_i (u_i^{(1)} - u_i^{(2)})$$

is identical with the expression $\sum_{i=1}^{q} \lambda_i d_i$ it follows that $\sigma_U^2 = D^2$, or a = 0.9173, and $a^{\frac{1}{2}} = 0.9578$. For the delinquent sample therefore, the mean of U turns out to be $\frac{1}{2}a$; for the non-delinquent sample the mean is $-\frac{1}{2}a$. The numerical value of these two statistics was found to be a = .4586 and -a = -.4586.

Solution of the Minimum Probability of Misclassification

By re-working Formulas 23 and 24 above, through the use of Formula 25, the probability of misclassifying an individual from either population is found to be .3161. Thus

TABLE III

CLASSIFICATION OF DELINQUENTS BY LEVEL OF DELINQUENCY USING LDF PROCEDURES FOR TWO RANDOMLY CHOSEN SAMPLES OF THE MAIN STUDY: SAMPLE A ON WHICH PREDICTION FORMULAE WERE DEVELOPED AND SAMPLE B ON WHICH THE FORMULAE WERE APPLIED THROUGH CROSS-VALIDATION

		Percentage of Correct Classification					
		Sample A (N = 175)		Sample B (N = 176)		Total (N = 351)	
Level*	Total	N	%	N	%	N	%
3	231	82	70.7	78	67.8	160	69.3
4	47	14.	60.9	18	75.0	32	68.1
5	31	7	46.7	12	75.0	19	61.3
6	22	9	81.8	8	73.0	17	77.3
7	12	4	66.7	5	83.3	9	75.0
8	8	4	100.0	3	75.0	7	87.5
Total	351	120	68.6	124	70.5	244	69.5

*In addition to the 351 delinquents in levels 3 to 8 above which had been randomly assigned to Samples A and B, 363 pupils had been classified as delinquents of levels 1 and 2. Application of the same procedures identified 75 or 42.3 percent of the level 1 group as delinquent and 111 or 60.0 percent of the level 2 group, an overall correct classification of 51.2 percent, approximately the same as chance expectation.

TABLE IV

A SUMMARY COMPARISON OF CLASSIFICATION EFFICIENCY FOR THE LDF PROCEDURES WITH CHANCE EXPECTATIONS FOR SAMPLES A' AND B' IN THE SCHOOL STUDY BY DELINQUENT AND NON-DELINQUENT GROUPS WITHIN THE SAMPLE

Samples Used		Correct Classification by Formula N	Correct Classification by Formula %	Correct Classification by Chance N	Correct Classification by Chance %	Percentage Improvement Over Chance
Sample A' - Delinquents	N = 91	60	65.9	45.5	50.0	15.9*
Sample A' - Non-delinquents	N = 91	72	79.1	45.5	50.0	29.1*
Total Sample A'	N = 182	132	72.5	91.0	50.0	22.5*
Sample B' - Delinquents	N = 91	63	69.2	45.5	50.0	19.2*
Sample B' - Non-delinquents	N = 91	63	69.2	45.5	50.0	19.2*
Total Sample B'	N = 182	126	69.2	91.0	50.0	19.2*
All Groups Total	N = 364	258	70.9	182.0	50.0	20.9*

*Significant at the 1 percent level

TABLE V

CLASSIFICATION OF DELINQUENTS BY LEVEL OF DELINQUENCY USING LDF PROCEDURES FOR TWO SAMPLES OF THE SCHOOL STUDY: SAMPLE A' ON WHICH PREDICTION FORMULAE WERE DEVELOPED AND SAMPLE B' ON WHICH THE FORMULAE WERE APPLIED THROUGH CROSS-VALIDATION

Level	Total	Sample A' (N = 91) N	Sample A' (N = 91) %	Sample B' (N = 91) N	Sample B' (N = 91) %	Total (N = 182) N	Total (N = 182) %
3	106	38	71.70	31	59.62	69	65.71
4	30	9	60.00	12	75.00	21	67.74
5	21	9	75.00	6	66.67	15	71.43
6	13	5	71.43	5	83.33	10	76.97
7	7	1	50.00	3	60.00	4	57.14
8	5	1	50.00	3	100.00	4	80.00
Total	182	63	69.23	60	65.93	123	67.58

$$p_1 = p_2 = .50$$
$$k = p_2/p_1 = 1.00$$
$$\log_e k = 0.00$$

and $\Pi_1 = \Pi_2 = p\left(z > \frac{0.0 - (-.4586)}{.9578}\right)$ or .3161

which is equal to 31.61 percent.

This means, in effect, that the theoretical probability of correct classification for both delinquents and non-delinquents is 68.39 percent. This overall percentage is essentially the same as that achieved by chance when the ratio of non-delinquents to delinquent in the population is known to be 80.5 to 19.5, as it is in this study. However, classification of delinquents by the LDF is more than three times as effective as chance procedures, the formula identifying 68.4 percent of the delinquents as against 19.5 percent identified by chance.

In actual practice, however, the population distribution is not known. In such a case, pure chance procedures make possible the identification of a maximum of 50.00 percent of each of the two populations.

Empirical Cross-Validation of Samples A and B

The actual effectiveness of the linear discriminant function was tested empirically by applying the formulas developed on Sample A to the delinquent and non-delinquent sub-samples in both Samples A and B. A study of cross-validation results as presented in Table II indicates that the overall percentage of correct classifications, namely 65.9, fell somewhat short of the predicted percentage of 68.4. The fact that the formula correctly identified only 58.9 percent of the non-delinquents in Sample B accounts for most of the shrinkage. However, 68.6 percent of delinquents in Sample A, and 70.5 percent in Sample B were correctly identified.

Classification results were also broken down in terms of delinquency level, as presented in Table III. The tabulation shows generally increasing proportions of correct classification with level of delinquency, level 5 being the only exception.

Results of the School Study

The set of linear equations composing the variance-covariance matrix for the six variables used in the school study is presented at the top of the next page.

The following is a list of the essential statistics derived:

$$LDF = .0219X_1 + .0282X_2 - .0430X_3 + .0398X_4 - .5247X_5 + .0564X_6$$
$$c = 2.7955$$

$$a = 1.2387;\ \sqrt{a} = 1.1129;\ \tfrac{1}{2}a = .6194$$

Using Formula 25 the probability of misclassifying an individual from either population is found to be equal to

$$\Pi_1 = \Pi_2 = p\left(Z > \frac{0.0 - (-.6194)}{1.1129}\right) =$$

$$p\,(z > .5566) = .2906$$

or approximately 29.1 percent, or a probability of correct classification for both delinquents and non-delinquents of 70.9 percent.

Results of cross-validation procedures (see Table IV) indicate that classification by the LDF permitted the correct identification of 65.9 percent of the delinquents in Sample A' and of 69.2 percent in Sample B'. Corresponding figures for non-delinquents are 79.1 percent and 69.2 percent for Samples A' and B' respectively. In each case improvement over chance is significant at the one percent level.

Although the two studies are not strictly comparable, the addition of school record data raised the overall predictive efficiency of the LDF from 65.9 percent in the main study to 70.9 percent in the school study, a difference significant at the 5 percent level.

Classification-results in terms of the level of delinquency are presented by means of Table V. Again, an increase in the proportion of correct classifications corresponds closely with increase in level of delinquency, a marked drop in predictive efficiency at level 7 being the only exception in the general trend.

Summary

This study has demonstrated the usefulness of applying multivariate statistical techniques to the analysis of MMPI scale scores alone, or in combination with school data, for the purpose of classifying ninth-grade boys as potential delinquents or non-delinquents. The techniques employed proved to be effective to the extent that 62.3 percent of the non-delinquents and 69.5 percent of the delinquents were correctly identified by the use of Multiphasic data alone, while a combination of Multiphasic and school record data made possible the correct identification of 74.2 percent of non-delinquent boys and 67.5 percent of delinquent-prone boys.

REFERENCES

1. Anderson, T. W. "Classification by Multivariate Analysis," Psychometrika, XVI (1951), pp. 31-50.
2. Collier, R. O. "Some Applications of the Method of Pivotal Condensation in Statistical Anal-

	Ma	Sc	Mf	Pd	HPR	Days	
Ma	125.8646 λ_1 +	42.2858 λ_2 +	8.5737 λ_3 +	26.3314 λ_4 -	.8156 λ_5 -	6.9509 λ_6 =	5.4505
Sc	42.2858 λ_1 +	107.1731 λ_2 +	19.2574 λ_3 +	37.0621 λ_4 -	.6524 λ_5 +	4.7766 λ_6 =	5.2088
Mf	8.5737 λ_1 +	19.2574 λ_2 +	80.9079 λ_3 +	2.5730 λ_4 +	.7312 λ_5 +	2.5101 λ_6 =	-2.8901
Pd	26.3314 λ_1 +	37.0621 λ_2 +	2.5730 λ_3 +	104.4952 λ_4 -	1.5153 λ_5 +	11.4942 λ_6 =	7.1099
HPR	- .8156 λ_1 -	.6524 λ_2 +	.7312 λ_3 -	1.5153 λ_4 +	.4401 λ_5 -	1.6429 λ_6 =	- .4515
Days	6.9509 λ_1 +	4.7766 λ_2 +	2.5101 λ_3 +	11.4942 λ_4 -	1.6429 λ_5 +	76.7578 λ_6 =	5.8242

ysis," Journal of Experimental Education, XXI (1953), pp. 233-41.

3. Hathaway, S.R., and Monachesi, E.D. Analyzing and Predicting Juvenile Delinquency with the MMPI (Minneapolis: University of Minnesota Press, 1953).
4. Johnson, P.O. .Statistical Methods in Research (New York: Prentice-Hall, Inc., 1949).
5. Lubin, A. "Linear and Nonlinear Discriminating Functions," British Journal of Psychology, Statistical Section, III (1950), pp. 90-103.
6. Mahalanobis, P.C., and others. "Anthropometric Survey of the United Provinces, 1941: A Statistical Study," Sankhya, IX (1939), pp. 89-324.
7. Rao, C. R. "The Problem of Classification and Distance Between Two Populations," Nature, CLX (1947), pp. 835-36.
8. Rao, C. R. "Tests of Significance in Multivariate Analysis," Biometrika, XXXV (1948), pp. 58-79.
9. Rao, C. R. Advanced Statistical Methods in Biometric Research (New York: John Wiley and Sons, 1952).

JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 28, Number 3, March 1960)

ON PROGRAMMING THE INTEGRAL RTH ROOT OF A NUMBER ON FRIDEN STANDARD AND SQUARE ROOT DESK CALCULATORS

WILLIAM J. MOONAN*
U. S. Naval Personnel Research Field Activity
San Diego, California

1. Introduction

IT IS THE purpose of this note to indicate some formulas which are useful for finding the roots of real numbers on Friden Calculators. The method used to derive the formulas is due to Newton although its application for taking advantage of the square-root feature on certain Friden Calculators is believed to be new. The mathematics of the derivations are given in the third section of the paper since this information is probably of least interest to those who will operate with the formulas. The second section contains three examples. The cube root example is symbolically programmed for the calculator in section 4. Table I presents a set of formulas which are efficient for finding the rth root of a real number. That is, for finding the square root, cube root, etc., of some number, N. The symbol x in the formulas stands for a guessed or trial value of the root and conceivably may be grossly in error. However, if it is considerably in error, the formulas will ultimately provide an accurate estimate of the root although it will take a longer time than if an estimate only slightly in error were used.

The method to use with the formulas involves essentially selecting a trial value, substituting this (as x) in the appropriate formula and thereby obtaining a new trial value. This new value is again substituted in the formula (as x). This iteration process is repeated until the desired decimal accuracy is obtained.

The formulas in Table I have been arranged so that the arithmetic ultimately involves the accumulation of two quotients. Therefore, the proper machine set-up requires the NON ENT key to be down. Also it is advisable that the ADD key is in the down position. In some of the formulas the divisors of x appear as improper fractions (e.g., $\frac{x}{4/3}$). In actual computations, these numbers can be changed to decimals appropriate to the accuracy required or the problem can be treated as a simple quotient (e.g., $3x/4$).

The Square Root Friden enables all roots of the form $r = 2^i$, where $i = 1, 2, 3, \ldots$, of a number to be taken directly without the use of any special formulas. For instance, to obtain the fourth root of 2.0000 we merely find the $\sqrt{2} = 1.41421$ and then find $\sqrt{1.41421} = 1.1892$ which is an accurate estimate of $2^{\frac{1}{4}}$. The higher order roots are difficult to obtain on the standard Friden since the procedure requires the trial values to be raised to large powers. The Square Root Friden is generally more efficient for both the high and low order roots since the quantities involved in the quotients are easier to calculate and moreover the square-root formulas provide estimates which converge faster to the root desired.

2. Examples

a) Consider first the problem of taking the cube root of 27. The answer, of course, is 3, but nevertheless let us see how the method works on a problem for which we know the exact answer. Suppose for our example that we wish our answer to be accurate to four decimal places. We set the decimal on the keyboard, lower dial and upper dial to one more than this, 5, and the upper dial decimal is also set to $2 \times 5 = 10$. Recall that the NON ENT and ADD keys are down. Since we will first use a standard Friden (or use the Square Root Friden in the standard manner) and we are seeking a cube root, the proper formula from Table I is:

$$\frac{N}{3x^2} + \frac{x}{1.5} .$$

We first guess a trial value of $x = 2.0000$, say,

*The opinions expressed are solely those of the author and are in no way official; nor are they to be construed as representing those of the U.S. Naval Personnel Research Field Activity or Bureau of Naval Personnel.

TABLE I

FORMULAS FOR FINDING ROOTS OF NUMBERS BY ITERATION

Root	Standard Friden	Square Root
2 (Square)	$\frac{N}{2x} + \frac{x}{2}$	$\sqrt{N} = N^{\frac{1}{2}}$
3 (Cube)	$\frac{N}{3x^2} + \frac{x}{1.5}$	$\frac{\sqrt{N}}{1.5\sqrt{x}} + \frac{x}{3}$
4 (Fourth)	$\frac{N}{4x^3} + \frac{x}{4/3}$	$N^{\frac{1}{4}}$
5 (Fifth)	$\frac{N}{5x^4} + \frac{x}{1.25}$	$\frac{N^{\frac{1}{4}}}{1.25x^{\frac{1}{4}}} + \frac{x}{5}$
6 (Sixth)	$\frac{N}{6x^5} + \frac{x}{1.2}$	$\frac{N^{\frac{1}{4}}}{1.5\sqrt{x}} + \frac{x}{3}$
7 (Seventh)	$\frac{N}{7x^6} + \frac{x}{7/6}$	$\frac{(xN)^{1/8}}{.875} + \frac{x}{7}$
8 (Eighth)	$\frac{N}{8x^7} + \frac{x}{8/7}$	$N^{1/8}$
9 (Ninth)	$\frac{N}{9x^8} + \frac{x}{1.125}$	$\frac{N^{1/8}}{1.25x^{1/8}} + \frac{x}{9}$
10 (Tenth)	$\frac{N}{10x^5} + \frac{x}{10/9}$	$N^{1/10}$

N = Number
x = Trial Value of Root

and copy this value on a worksheet. The number x is entered on the keyboard and also in the multiplier dial. The MULT key is punched and the resulting square is rounded to 5 decimal places and recorded on the keyboard. Then 4.00000 is multiplied by 3 and the product, 12.00000 is entered on the keyboard. The upper dial is cleared by hand. Next the number N = 27 is entered with the twirler knobs in the upper dial around the 10th decimal place and the dividing keys are pressed. Now 2.25000 appears in the lower dial which should now be locked. The upper dial is clear but in general it will not be, so that the CARR CLEAR will have to be punched. We have now found the first of the two quotients. Next x = 2.00000 is entered around the 10th decimal place in the upper dial and 1.50000 (the constant divisor) is entered on the keyboard. After dividing, we find 3.58333 in the lower dial. This number is our second approximation. After copying this on our worksheet we clear the upper dial and lower dial, and unlock the lower dial, and proceed to square 3.58333 and start repeating the evaluation of the formula. After six trials the worksheet would appear as:

Trial	x
1)	2.00000
2)	3.58333
3)	3.08980
4)	3.00257
5)	2.99999
6)	2.99999

Because the fifth and sixth trials produce the same number, we cease iterating since a stable trial value has been determined. Therefore, the estimate of the cube root of 27 is 3.0000 which is accurate to four decimal places.

b) The same problem will now be worked using the square-root feature. The appropriate formula from Table I, in this case, is:

$$\frac{\sqrt{N}}{1.5\sqrt{x}} + \frac{x}{3} \quad .$$

We first calculate directly and record the $\sqrt{N} = \sqrt{27} = 5.19615$ and $\sqrt{2} = 1.41421$ which is the square-root of our first trial value. Next 1.41421 is copied on the keyboard and multiplied by 1.5. This product, 2.12132, is copied on the keyboard. After clearing the upper dial by hand, 5.19615 is entered with the twirler knobs in the upper dial around the 19th decimal place. Division gives 2.44948 in the lower dial which should now be locked. Clear the upper dial. Then enter x = 2.00000 around the 10th place in the upper dial. Place 3.00000 (the constant divisor) on the keyboard and divide. The accumulated quotient is 3.11614 and this value is recorded on the worksheet as this is our second trial value. Next we calculate $\sqrt{3.11614}$ and repeat the process. The lower dial should be unlocked when starting a new iteration. The worksheet would appear as:

Trial	x	
1)	2.00000	$\sqrt{27} = 5.19615$
2)	3.11614	
3)	3.00108	
4)	3.00000	
5)	2.99999	

Notice that one less iteration is required for the same problem by this method than for the method used in sample a).

c) As a final example consider the problem of finding the fifth root of $\pi = 3.14159 = N$, using the Square Root Friden. From Table I the formula to be used is:

$$\frac{N^{\frac{1}{4}}}{1.25x^{\frac{1}{4}}} + \frac{x}{5}$$

The first task is to find $N^{\frac{1}{4}} = (3.14159)^{\frac{1}{4}}$. Since $\sqrt{3.14159} = 1.7724531$, then $N^{\frac{1}{4}} = \sqrt{1.7724531} = 1.33134$. This number is recorded on the worksheet. Because the fourth root of π is about 1.3, a good trial value for the fifth root is x = 1.20000. The fourth root of x is now found and multiplied by 1.25. The resulting product, 1.30830, is recorded on the keyboard and 1.33134 is entered around the 10th decimal place in the upper dial. Division gives 1.01761 and the lower dial is locked. Clear the upper dial, enter 1.20000 in the upper dial around the 10th decimal place, and record 5.00000 (the constant divisor) on the keyboard. Division produces 1.25761 in the lower dial and this value is the new x. After four trials, the worksheet appears as:

Trial	x	
1)	1.20000	$(3.14159)^{\frac{1}{4}} = 1.33134$
2)	1.25761	
3)	1.25726	
4)	1.25727	

Therefore, the fifth root of π is 1.2573 which is accurate to four decimal places.

3. Mathematical Derivations

Newton's method for finding roots of equations is well-known and is described in great detail in many theories of equations and calculus textbooks. Essentially the method says that the roots of an equation in x are approximated by the following expression:

$$x - \frac{f(x)}{f'(x)}$$

where f(x) is the function (or equation) of x, the f' (x) is its derivative. The functions of x used in this paper are very simple since we seek only the values of x satisfying the equation $x^r = N$, where r is some positive integer. Of course $f(x) = x^r - N$. We could consider more general functions but the result would be to introduce types of problems not frequently attempted. The derivative of f(x) is $f'(x) = rx^{r-1}$ so that:

$$x - \frac{f(x)}{f'(x)} = x - \frac{(x^r - N)}{rx^{r-1}} = \frac{N}{rx^{r-1}} + \frac{x}{r/(r-1)}$$

Upon substituting $r = 2, 3, \ldots, 10$ in the expression, we produce the formulas under the "Standard Friden" heading of Table I.

It occurred to the author that it may be more efficient to utilize the square-root feature on certain Friden Calculators to find the root of numbers. Accordingly, f(x) was modified to be:

$$f(x) = x^{r/2} - N^{\frac{1}{2}} .$$

Newton's expression leads to:

$$\frac{\sqrt{N}}{\frac{rx^{(r-2)/2}}{2}} + \frac{x}{r/(r-2)}$$

in this case. Substitution of r = 2 and 3 in the above expression gives the formulas for the square and cube roots in the second formula column of Table I. One could also use $r = 4, 5, \ldots,$ to generate iterative formulas for other roots, but it appears more wise to use $f(x) = x^{r/4} - N^{\frac{1}{4}}$ for r = 4, 5, and 6. This was done and the expression is:

$$\frac{N^{\frac{1}{4}}}{\frac{rx^{(r-4)/4}}{4}} + \frac{x}{r/(4-4)} .$$

For r = 7, 8 and 9, one might consider $f(x) = x^{1/8} - N^{1/8}$.

The general principle being used is now clear. The Square Root Friden formulas are derived from Newton's expression utilizing a function of the form $f(x) = x^{r/k} - N^{1/k}$ where $k = 2, 4, 8, 16, \ldots,$ Newton's formula for the general case is:

$$\frac{N^{1/k}}{\frac{rx^{(r-k)/k}}{k}} + \frac{x}{r/(r-k)} .$$

To find the rth root of N, locate r between two successive k's where k is an integral power of 2. In case r is equally distant between two k's use the smaller k. Substitute k and r in Newton's general formula above and solve for the specific iteration formula. Particular formulas are given in the third column of Table I.

4. Programming a Desk Calculator

The art of computation by means of desk calculators has been taught primarily by the demonstration of technique by the more experienced to the less. This unfortunate situation exists because it is usually much easier to show a student how to use a desk calculator than to tell him how, and adequate instructional materials are scarce. Calculator manufacturers have not been too successful in producing manuals which have been useful to novices. Frequently manuals and books on desk calculators have attempted to illustrate computations for particular problems. Consequently, it is often the reader's unpleasant duty to attempt to adapt the principles or steps of the computation (if they can be discovered) to the special task at hand.

Some of the difficulty could be overcome if an efficient code were devised for the instructions. While the author was engaged in writing statistical programs for a digital computer, it occurred to him that instructions similar to those used by computers could be devised for desk calculators. An extensive list of instructions for Friden calculators has been defined. Rather than list these instructions here, suffice it to say that they refer to a) the set-up of the registers and keys of the calculator, b) arithmetic operations, c) the transfer of information, d) branching and decision making, and e) tallying and control. These desk calculator instructions could be given a mnemonic code (such as CAD for "Clear and Add") in the same manner as digital computer instructions, but this does not seem particularly necessary and would impose an additional learning task. The advantage of desk calculator instructions include the increased generality of computation illustration, the compact description of computations, and a permanent and convenient record of procedures. The author believes that the actual demonstration of technique remains the most efficient method of learning desk calculator computation but he also believes the programming of often used and important procedures bears serious consideration, particularly for instruction manuals, and statistical and computation laboratories.

In order to illustrate how the programming might appear, a program for finding the cube root of N is provided in Table II. The reader is invited to follow simultaneously the verbal description in the first example in section 2 of this paper and the commands in Table II. Each command is given a number called a Step Number and a Reference. The Step Number provides the order in which the command is executed. The Reference refers to such things as a set of registers, a constant or variable, pair of numbers in registers, a

TABLE II

PROGRAM FOR COMPUTING THE CUBE ROOT OF N BY ITERATION ON A STANDARD FRIDEN CALCULATOR

Step No.	Command	Reference	Remarks
00	STANDARDIZE		Set calculator in standard condition.
01	DECIMAL	K;M;A;Q : d+1	Set registers to d+1 decimals.
02	DECIMAL	A : 2(d+1)	Set A register to 2(d+1) decimals.
03	NON ENTRY	DOWN :	Set Non Entry Key down.
04	ADD	DOWN :	Set Add Key down.
05	RECORD	x : d+1	Record approximate cube root, x, on data sheet.
06	KEYBOARD	x : d+1	Set Keyboard to x with d+1 decimals.
07	MULTIPLY	x : d+1	Multiply by x with d+1 decimals.
08	ROUND	A : d+1	Round product in A to d+1 decimals.
09	TRANSFER	A;K : d+1	Transfer A register to Keyboard.
10	MULTIPLY	3 : d+1	Multiply by 3.
11	ROUND	A : d+1	Round product in A to d+1 decimals.
12	TRANSFER	A;K : d+1	Transfer A register to Keyboard.
13	CLEAR	A, Q :	Clear A and Q registers by hand
14	SET	A;N : 2(d+1)	Set A register to N around 2(d+1) decimals. Use dial twirlers.
15	DIVIDE		Find quotient of A and K.
16	LOCK	Q :	Lock Quotient register.
17	CARR. CLEAR		Clear Carriage.
18	SET	A; x : 2(d+1)	Set A to x with 2(d+1) decimals.
19	KEYBOARD	1.5 : d+1	Set Keyboard to 1.5.
20	DIVIDE		Find Quotient of A and K.
21	COMPARE	Q; x : d	Compare Q register with last x to d decimals.
22	BRANCH, UNEQUAL	26 :	Go to instruction 26 if numbers compared are unequal.
23	ROUND	Q : d	If equal, round Q to d decimals.
24	RECORD	x : d	Record $x \doteq N^{1/3}$ on data sheet.
25	HALT		Stop calculations.
26	DEFINE	Q; x : d+1	Define numer in Q as the new trial x.
27	BRANCH	05 :	Go to instruction 05.

direction on how to set a register or key, or a step number. The Reference is divided essentially into two parts. The first part contains such things as those items just mentioned, and the second part (separated from the first by a colon) always specifies the decimal accuracy or decimal position if it is involved in the command.

The commands have not been specifically defined. By reading the "Remarks" column, the reader should have little difficulty in determining the proper interpretation. Before the reader considers Table II, he should know that the symbols used are defined as: K = Keyboard, M = Multiplier Dial, A = the Upper Dial or Register, Q = the Quotient or Lower Dial, d = the decimal accuracy desired, and x = a trial value of the cube root of N. The command "STANDARDIZE" means that all the dials and registers of the calculator are unlocked and cleared, and all the operational keys are put in their usual (UP) position.

JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 28, Number 3, March 1960)

MUST ALL TESTS BE MULTI-FACTOR BATTERIES?

EDWIN F. ROSINSKI
Medical College of Virginia
Richmond, Virginia

MUCH OF the present day literature on measurement concerns itself with the application and use of multi-factor test batteries. This trend has become so pronounced that journals are publishing series of articles on the use of these batteries.

In one of these series of articles, Super stated that if the detailed data on validity and norms in the Psychological Corporation's Differential Aptitude Tests and the United States Employment Service's General Aptitude Test Battery are as adequate as claimed, then:

> ....the use of singly developed and validated aptitude tests in counseling will in due course be virtually a thing of the past: the multipotentiality of the multifactor test battery gives too many advantages.[1*]

It appears as if Super's statement is not so much a prophecy but a recognition of a reality that is already present. With a pronouncement such as this and others[2] the ultimate question arises: Need single score tests be replaced by multi-factor batteries? It appears, however, that before this question can be answered, there is a need first to investigate the relationship between these single score tests and multi-factor batteries to ascertain if each is performing a different function or the same function. The Psychological Corporation's Differential Aptitude Tests (hereinafter referred to as the D. A. T.) and the World Book Company's Otis Quick Scoring Mental Ability Test: Gamma Form (hereinafter referred to as the Otis) were used to study this relationship. The decision to use these two tests appeared practical for two reasons: (1) the D.A.T. as a multi-factor test is gaining in prominence and use and includes in the manual a great deal of standardization data, and (2) the Otis is one of the most commonly used tests providing a single score IQ.

Procedure and Results

A sample of 153 students of a senior high-school class was selected. From the permanent record cards of these senior students the appropriate D. A. T. scores and the Otis IQ scores were taken. The former test had been administered to the students when they were in the tenth grade and the latter when the students were in the twelfth grade. There is no apparent reason to suspect that this recognizably selected sample should in any way contaminate the results. Using these raw data from the Otis and the D. A. T. scores, means, standard deviations and Pearson product-moment correlations were obtained.

After these correlations were found, it was necessary to determine partial r's and partial standard deviations.[3] Next the formula for Multiple R was used and a multiple correlation of .82 was computed. Not only does this prove to be significant at the .01 level of confidence but it is an extremely high correlation.

The next step involved the determination of the multiple regression equation or predicting the first variable (Otis) on the basis of the two known variables (D. A. T. Verbal and Abstract Reasoning Raw Scores). With the calculation of the partial regression coefficients, values were substituted in Cornell's[4] multiple regression or prediction equation and the Otis IQ score might then be predicted. This equation reads:

$$X_1 = b_{12.3} X_2 + b_{13.2} X_3 + a$$

where X_1 is the predicted Otis score; $b_{12.3}$ and $b_{13.2}$ the partial regression coefficients .51 and .89 respectively; X_2 the D. A. T. Verbal Reasoning raw score; X_3 the D. A. T. Abstract Reasoning raw score; and "a" the constant value in the equation determined to be 70. 5.

Therefore, with the above equation and known values, a hypothetical example can be tested. Assuming that student A has a D. A. T. Verbal and Abstract Reasoning scores of 29 and 34 respectively, what will his predicted Otis IQ be?

Substituting in the formula we have:

$$X_1 = b_{12.3} X_2 + b_{13.2} X_3 + a$$
$$X_1 = .51(29) + .89(34) + 70.5$$

*All footnotes will be found at end of article.

TABLE I

A PARTIAL TABLE FOR PREDICTING OTIS (GAMMA) IQ SCORES FROM D. A. T. VERBAL AND ABSTRACT REASONING RAW SCORES*

D. A. T. Abstract Reasoning Scores	D. A. T. Verbal Reasoning Scores											
	27	28	29	30	31	32	33	34	35	36	37	38
12	101	102	103	104	105	106	107	108	108	109	110	111
13	101	102	103	104	105	106	107	108	108	109	110	111
14	102	103	104	105	106	107	108	109	109	110	111	112
15	102	103	104	105	106	107	108	109	109	110	111	112

*The table is read in the following manner: If a D. A. T. Verbal Reasoning raw score is 34 and the Abstract Reasoning raw score is 14, the point at which the two columns intersect, the Otis IQ is read: 109.

TABLE II

ACTUAL AND PREDICTED OTIS (GAMMA) IQ SCORES WITH DIFFERENCES BETWEEN THE TWO

Student	Actual Otis IQ	Predicted Otis IQ	Difference	Student	Actual Otis IQ	Predicted Otis IQ	Difference
1	104	104	0	42	111	108	3
2	92	99	7	43	97	104	7
3	115	108	7	44	99	99	0
4	96	99	3	45	116	112	4
5	106	104	2	46	106	104	2
6	116	108	8	47	121	123	2
7	101	102	1	48	87	82	5
8	106	101	5	49	95	79	16
9	109	108	1	50	105	96	9
10	103	90	13	51	112	116	4
11	123	129	6	52	120	115	5
12	92	83	9	53	106	105	1
13	110	91	19	54	107	104	3
14	107	106	1	55	111	106	5
15	111	108	3	56	93	97	4
16	87	86	1	57	131	126	5
17	106	121	15	58	106	113	7
18	115	114	1	59	96	105	9
19	97	100	3	60	107	110	3
20	108	111	3	61	111	111	0
21	109	110	1	62	119	112	7
22	110	114	4	63	89	93	4
23	110	119	9	64	99	101	2
24	93	88	5	65	115	115	0
25	100	97	3	66	121	124	3
26	99	97	2	67	109	118	9
27	101	105	4	68	111	97	14
28	107	111	4	69	88	90	2
29	115	115	0	70	90	97	7
30	92	102	10	71	125	112	13
31	116	120	4	72	104	109	5
32	109	116	7	73	92	105	13
33	117	119	2	74	99	106	7
34	123	121	2	75	99	113	14
35	100	82	18	76	127	91	8
36	107	118	11	77	83	91	8
37	100	113	3	78	99	105	6
38	98	99	1	79	113	111	2
39	121	115	6	80	96	97	1
40	102	101	1	81	112	118	6
41	110	104	6	82	119	122	3

TABLE III

NUMBER AND PERCENT OF PREDICTIONS MADE WITHIN CERTAIN DIFFERENCES BETWEEN ACTUAL AND PREDICTED IQ'S

Diff. Between Actual and Predicted IQ	No. of Predictions	Percent of Predictions	Cumulative Percentage
0	5	7	7
1	10	12	19
2	19	11	30
3	11	13	43
4	8	10	53
5	8	10	63
6	5	7	70
7	8	10	80
8	2	2	82
9	5	7	89
10	1	1	90
11	1	1	91
12	0	0	91
13	3	3	94
14	2	2	96
15	1	1	97
16	1	1	98
17	0	0	98
18	1	1	99
19	1	1	100

TABLE IV

MEANS, STANDARD DEVIATIONS AND DIFFERENCES IN PREDICTED AND ACTUAL IQ SCORES

	Actual	Predicted	Difference
Mean	106.0	106.2	0.2
Standard Deviation	9.8	10.1	0.3

$X_1 = 14.79 + 30.26 + 70.5$
$X_1 = 115.55$

Rounded off to the nearest hundred, we have student A's predicted Otis IQ as 116.

Nevertheless, the task of predicting presented another problem. In attempting to predict several Otis IQ's it quickly became obvious that such a procedure of substituting in a formula was suitable with a small number of cases. On the other hand, if this was to be accomplished for a large group, it would be much too time consuming. Consequently, a prediction table was constructed by which the Otis score could be easily determined once the D. A. T. Verbal and Abstract Reasoning scores were known. Such a table was completed, but since it is large (49 by 47 columns) only a portion is here reproduced (Table I).

To test the prediction equation or, as it developed in this case, the prediction table, a new sample of students was selected. An entire senior class of 82 students in the same school where the original correlations were made was used. Again from the permanent record cards of these students the D. A. T. Verbal and Abstract Reasoning and the Otis scores were transcribed on a master work sheet. From the D. A. T. data the predicted IQ's were made and compared with the actual IQ's.

While the task of developing and recording the figures for the prediction table was tedious, the effort was rewarded by the ease with which the predicted scores were made.

Table II represents the findings of this procedure of predicting scores: the actual IQ scores along with the predicted IQ scores and the differences between the two. Prior to reading the table it is important to remember that one standard error of estimate for this prediction is 4.3 and the usual procedure is to expect, by the definition of a standard error of estimate, approximately 90% of the predictions within two standard errors of estimate, that is 90% of the predictions should fall within 8.6 (9) points of the predicted IQ. Table III shows the number and percent of predictions made within a certain difference.

From these two tables it is obvious that the predictions are not all perfect. This might well be expected; for perfect correlations, while highly desirable, are extremely unlikely. Yet, it does appear significant when considering the number of predictions that fell within the two standard errors of estimate and the reliability of the Otis itself. In this case, the prediction table and/or prediction equation appear to have some value.

Perhaps citing another figure might be more illustrative of the value of this prediction, for studies of such a nature as this are primarily concerned with mean scores and mean values. Therefore, when comparing the mean IQ's and standard deviations of the predicted and actual scores in Table IV, the accuracy of the predictions seems more vivid.

The greatest value, perhaps, of this study lies in this table. It appears almost extraordinary that predicting group mean IQ's could be accomplished with such a degree of accuracy as illustrated here.

Summary

In correlating the D. A. T. Verbal and Abstract Reasoning subtests with the Otis, a multiple correlation coefficient of .82 was found. From the multiple correlation coefficient, the prediction equation was developed and predicted 90% of the cases within two standard errors of estimate. While all the predictions were not perfect, the Multiple R was almost identical to the Otis reliability coefficient. In addition, while differences in individual predictions were noted, the predicted mean group IQ's were as close to the actual IQ's as possible (106 and 106.2).

Conclusions

From the data presented in this study, several conclusions seem appropriate.

1. Because of the significant and high Multiple R (.82) it appears as if the D.A.T. Verbal and Abstract Reasoning subtest are performing the same function as the Otis, i.e., measuring intelligence.
2. Because of the significant and high Multiple R, it appears as if the D.A.T. Verbal and Abstract Reasoning subtests can predict Otis IQ scores.
3. In actually predicting Otis IQ's on the basis of D.A.T. Verbal and Abstract Reasoning subtests, approximately 90% of the predictions will fall within nine points.
4. Predicting mean Otis IQ's for groups can be accomplished with considerable accuracy and confidence.

Implications

Several implications manifest themselves as a result of this study. If the multi-factor tests measure the same factors as some single score tests, then it only seems logical that they might ultimately replace these single score tests, but by the same token if they are measuring the same factors as single score tests, why not continue using these single score tests? It appears that even though the multi-factor tests can, with some degree of accuracy, do the same task as single tests the rush toward these multi-factor tests should be a considered one until the predictive value of these tests is substantiated with proven data. This is especially pertinent when considering the amount of time that each test takes. The Otis requires thirty minutes testing time while the D. A. T. requires approximately three hours, not to mention the time required for directions and scoring.

An argument for multi-factor test batteries could be the possible multipotentiality of the multi-factor tests, i.e., that the various subtests could be recombined to predict other phenomena.

What it seems to boil down to is this: that both tests tend to measure the same factors, but because they measure the same factors it does not necessarily follow that one should replace the other. Test users perhaps should consider some gross criteria before deciding on a particular test: (1) the purpose behind using a particular test, (2) the time allotted to accomplish the particular purpose, and (3) what the testor and testee are most comfortable in handling.

FOOTNOTES

1. Donald E. Super, "The Use of Multifactor Test Batteries in Guidance," Personnel and Guidance Journal, XXXV (September 1956), p. 14.
2. 1955 Invitational Conference on Testing Problems, sponsored by the Educational Testing Service of Princeton, New Jersey.
   Robert D. North, The Use of Multi-Factor Aptitude Tests in School Counseling Proceedings of the 1955 Invitation Conference on Testing Problems, New York, October 29, 1955, pp. 11-15.
3. Any standard test on statistics in psychology and education provides formulae for these values.
4. Francis G. Cornell, The Essentials of Educational Statistics (New York: John Wiley and Sons, 1956), p. 320.

JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 28, Number 3, March 1960)

AN EVALUATION OF THE LABORATORY IN A COLLEGE GENERAL BIOLOGY COURSE

DOUGLAS M. DEARDEN
University of Minnesota

THE RESEARCH reported to date in the literature indicates that no one type of laboratory is clearly superior in developing all of the outcomes desired from the science laboratory. The acceptable science studies indicate, for example, that the conventional type of laboratory has proven superior to other laboratory methods in teaching laboratory skills, but not in promoting the acquisition of facts.

The research reported in this study was conducted to help determine the contribution of several types of biology laboratories in a college biology course designed for freshmen and sophomores who do not plan to major in a science.

Experimental Treatments

The four experimental treatments compared were: (1) individual (conventional) laboratory, (2) demonstration laboratory, (3) workbook exercises, and (4) a biological term paper.

The individual laboratory work was based on the laboratory work found in many of the published college laboratory manuals, and involved a weekly two-hour laboratory period. The demonstration laboratory required one laboratory hour per week. The laboratory instructors had set up experiments and materials prior to the laboratory period for the students to observe and study. This demonstration type of laboratory eliminated many of the more routine aspects of the laboratory work for the student and allowed them to concentrate on the important points and principles.

The workbook exercises were submitted weekly and were based on questions in the demonstration laboratory manual. The answers to these exercises had to be obtained from non-laboratory sources, such as texts. The biological report was used as an experimental control.

The lecture portion of the class was presented by one lecturer who kept the lectures as constant as possible for all sections of the class.

Measuring Instruments

Three tests were used to evaluate the achievement of students randomly assigned to the four experimental treatments. The students took these tests before beginning the course, upon completion of the course, and three months after completing the course. The three tests were tests of Biological Knowledge, Scientific Thinking, and Biological Attitudes. These tests were validated and the reliability coefficients for the tests were computed. The reliability coefficients were computed by the split-halves method, using the maximum likelihood estimate of the correlation coefficient and then correcting this coefficient with the Spearman-Brown formula. The corrected reliability coefficients were 0.87 for the Biological test, 0.86 for the Biological Attitude test and 0.85 for the Scientific Thinking test.

Experimental Population

The students involved in the experiment were the 924 students who completed the one quarter General College Human Biology course, G.C. 10-A during the 1958-59 academic year. These students were enrolled in the General College, Home Economics, Education, Mortuary Science, and the College of Science, Literature, and the Arts. The students in the six sections of this course were randomly assigned in approximately equal numbers to the treatments through the use of a table of random numbers.

Experimental Design

A stratified randomized design was used in this study, with the experimental treatments being randomly assigned over all of the experimental material. The students were stratified for college and sex prior to being assigned to treatments as a means of more equally distributing the students of various abilities.

The desirable statistical tool to test the null hypotheses of the study was the analysis of variance. Of the two main models of this analysis, the fixed effects model, Model I, is designed for studies such as the present study. T-tests, Duncan's Multiple Range Test, and the Johnson-Neyman Technique were also included in the statistical analyses.

Experimental Results

A comparison of the initial and final test scores

by means of t-tests indicated that each of the four treatments had significantly increased student's scores on the three tests. The treatment which most consistently increased the variance on the final tests was the individual laboratory, which suggests that this treatment allowed more adequately for individual differences in student academic ability.

The overall results from the experiment are presented in Tables I, II, and III which give the means of the tests given at the end of the experimental period for each of the six sections.

When the final Biological Knowledge Test scores were analyzed by a two-way analysis of variance (treatment and ability level) none of the treatments was signifcantly better in increasing student scores. Table IV presents the analysis of variance results for the lecture section which most closely approached significant results between treatments on this test.

The conclusion which resulted from the analysis of the Biological Knowledge test was that none of the four treatments was significantly better in helping students learn the factual aspects of biology.

The Scientific Thinking tests were similarly analyzed by the analysis of variance technique to determine which treatment functioned best in improving scientific thinking by students. The differences in achievement on this test were not statistically significant, indicating that none of the treatments was superior in improving scientific thinking by students.

When the Biological Attitude Test results were analyzed, significant differences were found between the treatment groups in two of the lecture sections. The significant differences in section one, Fall quarter (Table V) were further analyzed by Duncan's Multiple Range Test, which indicated that the individual laboratory group of this section had significantly higher attitude test scores than the other three treatments. In section three, Fall quarter, the workbook group was significantly higher on this test than the other three treatments. However, the total results do not indicate that any one treatment was consistently superior in improving student's scores on this attitude test.

The preceeding analysis of variance results also were used to determine if any treatment was superior so far as being more suitable for students of comparatively high, middle, or low academic ability. The chosen criterion of academic ability was the student's ACE scores. The analysis of variance results indicated that the students of a given ability group achieved equally well on the tests no matter which treatment they were assigned to. Thus none of the treatments proved to be better for the students of any of these three comparative levels of ability.

Analysis of the Follow-up Tests

Three months after completing the experimental course, a random sample of students again took the tests used in this study. Of the 122 students selected, 79.5 percent took the three tests. T-tests for correlated measures were used to compare the final and follow-up test means and variances.

The students in all treatment groups tended to score lower on the follow-up tests of Biological Knowledge and Attitudes. However, each of the treatment groups increased their scores on the follow-up Scientific Thinking Test. In comparing the variances of the final and follow-up tests, the individual laboratory group most consistently increased the variability in scores on the follow-up tests, indicating that this treatment allowed more adequately for individual differences in learning.

A two-way analysis of variance of the follow-up tests (treatments × number of additional biology courses taken by students) indicated that the four treatments were not significantly different in promoting a more permanent learning by students of biological attitudes, facts, or scientific thinking. This conclusion held both for those students who did or did not take an additional biology course after completing the experimental course.

Johnson-Neyman Technique*

While no significant differences were found between the treatments on the Biological Knowledge and Scientific Thinking Tests, some differences were noted on the Biological Attitude Test. Thus it was advisable to collate the information from all of the six single experiments to see if the combined data indicated any significant differences between the treatments. A qualitative modification of the Johnson-Neyman Technique was used to combine and collate the data from the six separate experiments.

In this analysis, two treatments are compared at one time. For example, when the combined demonstration laboratory groups were compared with the combined individual laboratory groups on the Attitude test, the difference (D) in mean final Attitude test scores between the two treatments and the variance (S_D^2) of D were computed. Then the value of t_o was calculated, where

$$t_o = \frac{D}{S_D^2} .$$

Referring to the table of t, with the appropriate degrees of freedom, the null hypothesis of no mean differences between the two groups could be re-

* The technical background for this technique is presented in the text by Palmer O. Johnson and Robert Jackson, Modern Statistical Methods: Descriptive and Inductive (Chicago: Rand McNally and Co., 1959), 427-8.

TABLE I

MEAN SCORES OF STUDENTS ON THE FINAL BIOLOGICAL KNOWLEDGE TEST

	Lecture Section					
Treatment	Sec. 1 Fall	Sec. 2 Fall	Sec. 3 Fall	Sec. 1 Winter	Sec. 2 Winter	Sec. 1 Spring
Demonstration Lab.	54.76	50.04	52.92	50.64	55.50	48.81
Workbook	52.31	52.06	55.33	51.29	53.36	53.05
Report	51.44	49.75	53.67	49.69	52.60	50.42
Individual Lab.	56.40	53.13	56.42	46.13	52.83	50.62

TABLE II

MEAN SCORES OF STUDENTS ON THE FINAL SCIENTIFIC THINKING TEST

	Lecture Section					
Treatment	Sec. 1 Fall	Sec. 2 Fall	Sec. 3 Fall	Sec. 1 Winter	Sec. 2 Winter	Sec. 1 Spring
Demonstration Lab.	32.04	30.20	28.82	30.73	32.97	34.52
Workbook	32.19	31.50	31.33	32.26	33.41	31.62
Report	30.75	31.78	29.84	30.00	32.29	32.58
Individual Lab.	35.53	30.00	29.75	28.33	29.50	35.90

TABLE III

MEAN SCORES OF STUDENTS ON THE FINAL BIOLOGICAL ATTITUDE TEST

	Lecture Section					
Treatment	Sec. 1 Fall	Sec. 2 Fall	Sec. 3 Fall	Sec. 1 Winter	Sec. 2 Winter	Sec. 1 Spring
Demonstration Lab.	11.27	11.58	11.82	13.73	12.81	12.57
Workbook	12.35	12.56	14.15	13.76	13.23	13.19
Report	12.00	9.90	13.86	12.38	14.07	15.46
Individual Lab.	15.87	14.47	12.75	12.60	13.22	13.19

TABLE IV

ANALYSIS OF VARIANCE OF THE BIOLOGICAL KNOWLEDGE TEST
(Section One Fall Quarter)

Source of Variation	D. F.	Sum of Squares	Mean Square	F-Test
Between Ability Levels (i)	2	4169.8974	2084.9487	$F = 16.32$ $P < .05$ Reject the Ho of equal means among ability levels.
Between Treatments (j)	3	448.8613	149.6204	$F = 1.17$ $P > .05$ Do not reject the Ho of equal treatment means.
Interaction (i) × (j)	6	1651.5720	275.2620	$F = 2.15$ $P > .05$ Do not reject the Ho of no interaction.
Error	144	18390.5668	127.7123	
Total	155	24660.8975		

TABLE V

ANALYSIS OF VARIANCE OF THE ATTITUDE TEST
(Section One Fall Quarter)

Source of Variation	D. F.	Sum of Squares	Mean Square	F-Test
Between Ability Levels (i)	2	46.0385	23.0193	$F = 0.88$ $P > .05$ Do not reject the H_O of equal means among the ability levels.
Between Treatments (j)	3	243.1798	81.0599	$F = 3.10$ $P < .05$ Reject H_O of equal treatment means.
Interaction (i × j)	6	82.0198	13.6700	$F = 0.52$ $P > .05$ Do not reject the H_O of no interaction.
Error	144	3763.4543	26.1351	
Total	155	4134.6924		

TABLE VI

COMPARISON OF EXPERIMENTAL TREATMENTS USING THE JOHNSON-NEYMAN TECHNIQUE

Test*	Treatments Compared	D. F.	D	S_D	t_o	Probability of t_o
A	Report - Individual Lab.	374	-0.72	0.70	1.02	$.3 < P < .4$
	Workbook - Individual Lab.	371	-0.48	0.64	0.75	$.4 < P < .5$
	Workbook - Report	536	0.53	0.50	1.06	$.2 < P < .3$
	Demonstration Lab. - Individual Lab.	372	-1.09	0.64	1.70	$.05 < P < .1$
	Demonstration Lab. - Report	537	-0.20	0.50	0.41	$.6 < P < .7$
	Demonstration Lab. - Workbook	534	-0.73	0.47	1.55	$.1 < P < .2$
B	Report - Individual Lab.	374	-0.03	0.97	0.03	$.9 < P < 1.0$
	Workbook - Individual Lab.	370	0.31	0.96	0.33	$.7 < P < .8$
	Workbook - Report	535	0.03	0.72	0.04	$.9 < P < 1.0$
	Demonstration Lab. - Individual Lab.	372	0.42	1.00	0.41	$.6 < P < .7$
	Demonstration Lab. - Report	537	-0.36	0.74	0.49	$.6 < P < .7$
	Demonstration Lab. - Workbbok	533	-0.54	0.73	0.73	$.4 < P < .5$
S	Report - Individual Lab.	374	0.69	1.13	0.60	$.5 < P < .6$
	Workbook - Individual Lab.	371	0.66	1.10	0.60	$.5 < P < .6$
	Workbook - Report	536	0.37	0.83	0.44	$.6 < P < .7$
	Demonstration Lab. - Individual Lab.	372	0.47	1.08	0.44	$.6 < P < .7$
	Demonstration Lab. - Report	537	-0.07	0.82	0.09	$.9 < P < 1.0$
	Demonstration Lab. - Workbook	534	-0.47	0.80	0.59	$.5 < P < .6$

* A = Attitude final test
B = Biological knowledge final test
S = Scientific thinking final test

jected or accepted.

Using this technique to compare each of the combined treatments with every other combined treatment, the conclusion was drawn that none of the mean test scores were significantly higher at the .05 level of significance. Thus, at this level, none of the treatments was significantly better in improving the mean final scores of students on these three tests. However, at the .10 level, the individual laboratory, when compared with the demonstration laboratory, was associated with a significant increase in scores on the Biological Attitude Test. A summary of the results of all of these comparisons is presented in Table VI.

A sample of the experimental gr ps of students was compared with two random samples of students who had completed this course several years ago. The ACE scores of these samples of students were compared by an analysis of variance technique. The experimental group was not significantly different, with regard to ACE scores, from the students of previous years. Thus the results of this experiment can be generalized to that population of students taking the experimental course, assuming that the characteristic of future students do not change.

Summary

Four experimental treatments (individual or conventional laboratory, demonstration laboratory, workbook exercises, and a biological term paper) were statistically evaluated in a college general biology course. Tests of Biological Knowledge, Scientific Thinking, and Biological Attitudes were used to compare the achievement of students randomly assigned to these treatments.

A significant increase in test scores was associated with each treatment at the end of the instructional period, but none of the treatments was superior to the other treatments in increasing student's scores. Each of the treatments allowed for individual differences in the academic ability of students, but the individual laboratory allowed more adequately for these differences in learning ability and thus allowed the students to learn at a rate more consistent with their ability.

In comparing the test results of students with comparatively high, middle, and low academic ability, none of the treatments was found to be more suited to students of a particular level of ability.

At the .10 level, the Johnson-Neyman Technique indicated a difference between two of the treatments. When the individual laboratory and demonstration laboratory groups were compared, the individual laboratory groups had significantly higher mean test scores on the Biological Attitude Test.

None of the treatments was significantly superior in promoting a more permanent learning of the course materials, as measured by tests administered three months after the students had completed the experimental course.

The results of experimentation can provide useful practical results. Based on the results of this experiment, a combination of these methods could be used to good advantage. The advantages of all four treatments could be incorporated into a course, with particular emphasis being placed on the individual laboratory because of its better adjustment to individual differences in learning ability of students. In addition, if the four treatments were carefully integrated into a course, student interest in the course materials could be increased.

JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 28, Number 3, March 1960)

APPLICATION OF MATRICES IN THE ANALYSIS OF SOCIOMETRIC DATA

JOHN X. JAMRICH
Michigan State University
East Lansing, Michigan

A NUMBER OF research workers in the field of sociometry have clearly identified the inadequacy of the ordinary methods used in the analysis of sociometric data. Briefly, the inadequacies stem from the fact that ordinary methods of analysis do not determine, uniquely, status levels of the individuals or of sub-groups of individuals in the group being studied. Furthermore, present methods lack specific quantitative characteristics to permit analysis of change in the social structure of a group, in the relative status of an individual within the group, or of the sub-groups, usually referred to as cliques.

Katz* has presented a mathematical derivation of a "status vector" of a given group. The derivation of the numerical elements of this vector is a relatively cumbersome process, involving the solution of "n" simultaneous equations, where "n" is the number of individuals comprising the group.

It is the purpose of this paper to present the development of a relatively simple numerical characteristic for sociometric data which will permit:

1. determination of the status of an individual within the group, taking into account his choices and the kinds of choices by others in which he is involved;

2. determination of a characteristic vector for the entire social group so that it may be compared with other groups and with sociometric data for the same group at a later date;

3. the identification of cliques, and their relative status within the group, and changes which may be taking place from time to time as new sociometric data are gathered for the same group.

The understanding of the development and computation is based on the simple concept of addition and multiplication of matrices. By definition, the sum of two matrices:

$$(1)\quad M_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ and } M_2 = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

is given by

$$(2)\quad M_1 + M_2 = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}$$

The product of two matrices, with elements as above, is defined as:

$$(3)\quad M_1 \cdot M_2 = \begin{pmatrix} a_{11}\,b_{11} + a_{12}\,b_{21} & a_{11}\,b_{12} + a_{12}\,b_{22} \\ a_{21}\,b_{11} + a_{22}\,b_{21} & a_{21}\,b_{12} + a_{22}\,b_{22} \end{pmatrix}$$

Thus, if we have a matrix

$$(3a)\quad M_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

then, M_1^2 would be M_1 multiplied by itself, namely:

$$(4)\quad M_1^2 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

and

$$(5)\quad M_1 + M_1^2 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}$$

For later references, it should be pointed out

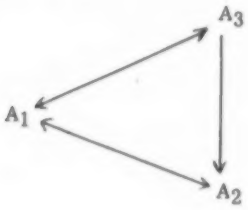
*Leo Katz. "A New Status Index Derived from Sociometric Analysis," Psychometrika, XVIII (March 1953), pp. 39-43.

that the sums of the columns of the matrix $M_1 + M_1^2$ are

(6) $(5, 5, 3)$.

Consider, as an example, the following simple sociogram involving three individuals, A_1, A_2, and A_3, each individual having made one or two choices of equal "strength."

(7)



The matrix representing this sociogram will be:

(8)
$$M_1 = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \end{array} \begin{array}{c} \begin{array}{ccc} A_1 & A_2 & A_3 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \end{array}$$

In the above matrix (8), the entries in the columns indicate that A_1 is chosen by certain of the other individuals, and the entries in the rows indicate that A_1 has chosen certain of the other people in the group. In other words, the entries represent the number of one-stage relations (or communications) between the individuals.

The above matrix (8), is seen to be precisely the one used as an illustration in (3a). Now, let us multiply the matrix in (8) by itself. We obtain:

(9)
$$M_1^2 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

If we consider M_1 to be a communication matrix, then in the square of it, M_1^2, the entry in the ith row, jth column gives the number of "2-stage" communications between A_i and A_j. For example, A_3 can communicate once with A_1 in two stages, while A_2 cannot communicate with A_1 in two stages. We might say, then, that the entries represent the number of two-stage relations among the corresponding individuals.

Now forming the sum, $M_1 + M_1^2$, we obtain the matrix (10) shown at the top of column two on this page.

From the interpretation of (8) and (9), we see that the entries in (10) represent the number of

(10)
$$M_1 + M_1^2 = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}$$

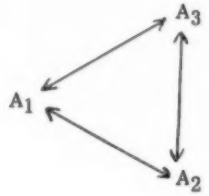
The column sums of (10) are $(5, 5, 3)$

one- and two-stage relationships between corresponding individuals. Hence, the column sums will give the total number of one- and two-stage "inward" relations (including the one with himself) enjoyed by each individual. This latter fact is used here as an indicator of the status of the individual within the group. (The row sums can be interpreted as "outward" one- and two-stage relations or communications.)

Now, a consideration of the matrix in (10) and the original sociogram reveals some interesting facts. First, the diagonal of the matrix in (10) gives, for each of the individuals, the number of "mutual choices" enjoyed by each one. Thus, A_1 has two mutual choices, while A_2 and A_3 have only one each. As indicated above, the column sum vector $(5, 5, 3)$, gives the relative status of each individual within the group. Thus, according to this characteristic vector, A_1 and A_2 have equal status, while A_3 has less than either of the others.

Using this same group of three individuals, let us illustrate the manner in which the column-sum vector can be used to compare change in a given group from one time to another. For this purpose, let us suppose that a subsequent application of the same procedure yielded the following sociogram:

(11)



whose matrix is,

$$M_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

The square of this matrix is

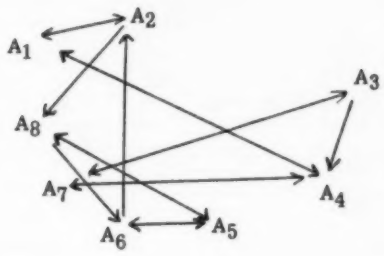
(12)
$$M_1^2 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

and the sum,

$$M_1 + M_1^2 = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$$

The column-sum vector is (6, 6, 6). In other words, compared with the column-sum vector of the sociogram of (7), it is evident that all three of the individuals enjoy two mutual choices, as revealed by the diagonal values in (12), and, they are all of equal status.

Remembering that these illustrations assume two choices for each individual, and two choices of equal strength, let us illustrate the utilization of the method with a more complex group, say, of eight individuals. First, let us consider the following sociogram for these eight:



The matrix for this sociogram would be

$$(13) \qquad C = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

and the square of it is

$$(14) \qquad C^2 = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 2 \\ 1 & 0 & 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

The sum, $C + C^2$ is shown at the top of the next column, and the resulting column-sum vector is

(16) V: (7, 6, 3, 8, 6, 6, 6, 6)
$(A_1\ A_2\ A_3\ A_4\ A_5\ A_6\ A_7\ A_8)$

From the values of the diagonal in (14), it is clear that A_1, A_4, A_5, and A_7 each have two mutual choices to their credit. From the characteristic column-sum vector, it is seen that A_4 has the "most" status, which fact can be substantiated from the original sociogram, S_1, for the purposes of this discussion. Similarly, A_1 is seen to have the next highest status designation, followed by A_5, A_6, A_7 and A_8. The latter four are indicated by the vector sums as being of equal status, a fact which can again be checked in the original sociogram.

$$(15) \qquad C + C^2 = \begin{pmatrix} 2 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 & 0 & 0 & 2 & 0 \\ 1 & 1 & 1 & 2 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 2 & 2 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 2 \\ 1 & 0 & 1 & 2 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 & 2 & 2 & 0 & 1 \end{pmatrix}$$

It may be worth noting that the column-sum vector values of A_1 and A_5 are 7 and 6 respectively. From the sociogram, one would conclude that these two individuals have equal status on the basis of the fact that each enjoys two mutual choices. The values of 7 and 6 in the column-sum vector indicate the fact that A_1 has one more two-stage "inward" relation in the group than does A_5, a fact which may be of some importance in the inter-group relationships.

The interpretation of the column sums might be illustrated at this point by showing the meaning of the number "7" which is the value associated with A_1. The seven "inward" one- and two-stage relations for A_1 are as follows:

$A_1 \leftarrow A_2 \leftarrow A_6$

$A_1 \leftarrow A_2 \leftarrow A_1$

$A_1 \leftarrow A_4 \leftarrow A_1$

$A_1 \leftarrow A_4 \leftarrow A_3$

$A_1 \leftarrow A_4 \leftarrow A_7$

$A_1 \leftarrow A_2$

$A_1 \leftarrow A_4$

By using the characteristic column-sum vector to identify the individuals with the most status, and using the original matrix, we are able to identify the cliques in this group and associate with them a status designation by using the column-sum vector for each individual in the clique.

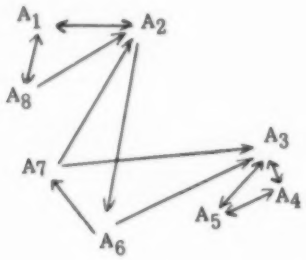
For example, from (16) we see that A_4 has the highest column sum. From the diagonal entry in (14) we see that A_4 enjoys two mutual choices which, from (13) are with A_1 and A_7. Hence, A_4, A_1, and A_7 can be said to form a clique.

Similar analyses yield the following cliques,

the first three of which actually constitute a large sub-group:

$$(17)\qquad \begin{array}{llll} A_4 & A_1 & A_7 & : (8, 7, 6) \\ A_4 & A_3 & A_7 & : (8, 3, 6) \\ A_1 & A_2 & A_4 & : (7, 6, 8) \\ A_5 & A_6 & A_8 & : (6, 6, 6) \end{array}$$

Now consider the same eight individuals with the following sociogram at some later time.



The matrix representing this sociogram is

$$(18)\qquad D = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The square of this matrix is

$$(19)\qquad D^2 = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

and

$$(20)\qquad D + D^2 = \begin{pmatrix} 2 & 2 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 2 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

The characteristic column-vector for (20) is

$$(21)\qquad \begin{array}{cccccccc} (6, & 7, & 10, & 8, & 8, & 4, & 2, & 3\ ) \\ (A_1, & A_2, & A_3, & A_4, & A_5, & A_6, & A_7, & A_8) \end{array}$$

and the cliques, with their characteristic column-sum vectors, are

$$\begin{array}{lll} A_1\ A_2\ A_8 & : & (6,\ 7,\ 3) \\ A_3\ A_4\ A_5 & : & (10,\ 8,\ 8) \end{array}$$

It is evident that the individual with the most status is A_3, quite a change from the case in (13). The status of the other individuals can be read quite easily from (21). There are two strong cliques, the stronger of them being A_3, A_4, A_5.

Summary

The use of the characteristic column-vector presented above appears to provide a useful method for quantifying sociometric data. Specifically, it affords a relatively simple procedure for determining a numerical value for the status of individuals in a group.

It should be emphasized, however, that the method presented above utilizes the notion of the incoming, one- and two-stage relations enjoyed by each individual in the group. This procedure, then, places additional emphasis not only upon how many chose a given individual, but whether those who chose him were chosen by others in the group.

Finally, the utility of the method of the paper will depend upon additional research with sociometric data to ascertain the significance of the two-stage relationships in interpersonal choices of individuals.

JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 28, Number 3, March 1960)

A TABLE TO FACILITATE THE CALCULATION OF NORMALIZED T-SCORES FOR SMALL SAMPLES (N>5<51)

PETER F. MERENDA
Walter V. Clarke Associates, Inc.
East Providence, Rhode Island

SEVERAL AUTHORS of psychological measurement and statistics texts, Garrett (1) and Edwards (2), for example, have included in the appendix a table to facilitate the calculation of T-scores. The use of this table, however, requires the performing of all of the steps in the usual procedure for normalizing distributions and calculating T-scores.

The table presented in this article requires only that the first four steps in the usual procedure be performed. The final step of calculating percentages, which is the most time-consuming of all the steps, before entering the T-score table is eliminated. In order to enter this table, all one needs to know are the number in the sample and cumulative frequency below a given score plus one half of the frequency of the given score. The table is divided into two parts: Table I for N's between 6 and 24 inclusive, and Table II for N's of 25-50 inclusive. The procedure and instructions for use with this table are given below with example:

INSTRUCTIONS FOR USE WITH T-SCORE TABLE

Step	Procedure
1.	Prepare frequency distribution of scores.
2.	Cumulate scores, i.e., add each successive row of scores.
3.	Tabulate the frequency below a score + 1/2 the frequency of a given score.
4.	Enter T-score table with N (number in sample) and read standard normalized scores in the appropriate cell for values calculated in step 3.

Example:

N = 10

X	f	Cum. f.	Cum. f. below +1/2 f	T-scores (Read from chart)
100	1	10	9.50	67
90	2	9	8.00	58
85	3	7	5.50	51
80	2	4	3.00	45
75	1	2	1.50	40
70	1	1	.50	33

TABLE I

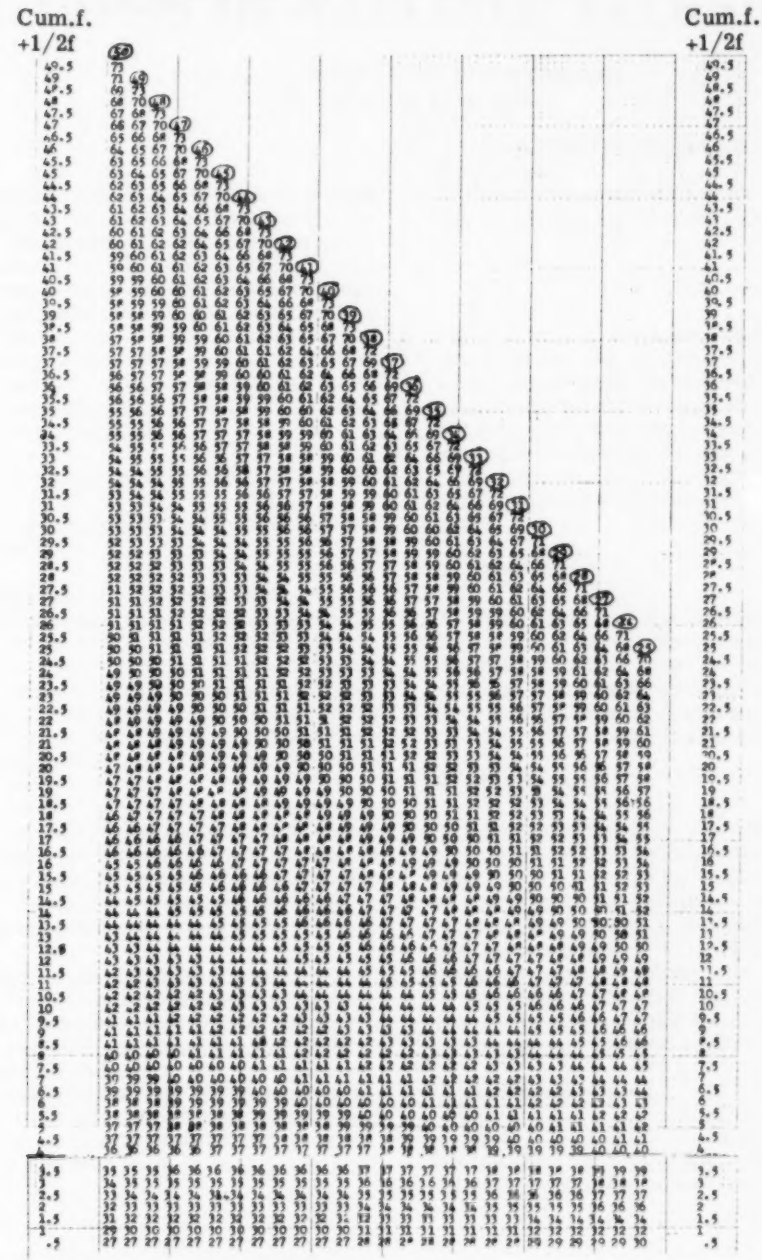
A TABLE FOR CALCULATING T-SCORES
(N > 5 < 25)

Cum.f. +1/2f	N 24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	Cum.f. +1/2f
23.5	70															
23	68	23 70														
22.5	66	67	22 70													
22	64	65	67	21 70												
21.5	63	63	65	67	20 70											
21	62	62	63	65	67	19 70										
20.5	61	61	62	64	66	68	18 69									
20	60	60	61	63	65	66	68	17 69								
19.5	59	59	60	62	64	65	67	68	16							
19	58	58	59	61	62	64	65	66	67	15 68						
18.5	57	57	58	60	61	62	64	65	66	67	14 68					
18	57	56	57	59	60	61	62	63	65	66	67	13 68				
17.5	56	55	56	58	59	60	61	62	63	64	65	67	12 67			
17	55	54	55	57	58	59	60	61	62	63	64	65	66	11 67		
16.5	54	53	54	56	57	58	59	60	61	62	63	64	65	66	10 67	
16	53	52	53	55	56	57	58	59	60	61	62	63	64	65	66	9.5
15.5	52	51	52	54	55	56	57	58	59	60	61	62	63	64	65	9
15	51	50	51	53	54	55	56	57	58	59	60	61	62	63	64	8.5
14.5	50	49	50	52	53	54	55	56	57	58	59	60	61	62	63	8
14	49	48	49	51	52	53	54	55	56	57	58	59	60	61	62	7.5
13.5	48	47	48	50	51	52	53	54	55	56	57	58	59	60	61	7
13	47	46	47	49	50	51	52	53	54	55	56	57	58	59	60	6.5
12.5	46	45	46	48	49	50	51	52	53	54	55	56	57	58	59	6
12	45	44	45	47	48	49	50	51	52	53	54	55	56	57	58	5.5
11.5	44	43	44	46	47	48	49	50	51	52	53	54	55	56	57	5
11	43	42	43	45	46	47	48	49	50	51	52	53	54	55	56	4.5
10.5	42	41	42	44	45	46	47	48	49	50	51	52	53	54	55	4
10	41	40	41	43	44	45	46	47	48	49	50	51	52	53	54	3.5
9.5	40	39	40	42	43	44	45	46	47	48	49	50	51	52	53	3
9	39	38	39	41	42	43	44	45	46	47	48	49	50	51	52	2.5
8.5	38	37	38	40	41	42	43	44	45	46	47	48	49	50	51	2
8	37	36	37	39	40	41	42	43	44	45	46	47	48	49	50	1.5
7.5	36	35	36	38	39	40	41	42	43	44	45	46	47	48	49	1
7	35	34	35	37	38	39	40	41	42	43	44	45	46	47	48	.5
6.5	34	33	34	36	37	38	39	40	41	42	43	44	45	46	47	
6	33	32	33	35	36	37	38	39	40	41	42	43	44	45	46	
5.5	32	31	32	34	35	36	37	38	39	40	41	42	43	44	45	
5	31	30	31	33	34	35	36	37	38	39	40	41	42	43	44	
4.5	30	29	30	32	33	34	35	36	37	38	39	40	41	42	43	
4	29	28	29	31	32	33	34	35	36	37	38	39	40	41	42	
3.5	28	27	28	30	31	32	33	34	35	36	37	38	39	40	41	
3	27	26	27	29	30	31	32	33	34	35	36	37	38	39	40	
2.5	26	25	26	28	29	30	31	32	33	34	35	36	37	38	39	
2	25	24	25	27	28	29	30	31	32	33	34	35	36	37	38	
1.5	24	23	24	26	27	28	29	30	31	32	33	34	35	36	37	
1	23	22	23	25	26	27	28	29	30	31	32	33	34	35	36	
.5	22	21	22	24	25	26	27	28	29	30	31	32	33	34	35	

Cum.f. +1/2f	N 6	7	8	9	Cum.f. +1/2f
0.5	36	35	35	34	0.5
1.0	40	39	39	38	1.0
1.5	43	42	41	40	1.5
2.0	46	44	43	42	2.0
2.5	48	46	45	44	2.5
3.0	50	48	47	46	3.0
3.5	52	50	49	47	3.5
4.0	54	52	50	49	4.0
4.5	57	54	51	50	4.5
5.0	60	56	53	51	5.0
5.5	64	58	55	53	5.5
		61	57	54	6.0
		65	59	56	6.5
			61	58	7.0
			65	60	7.5
				62	8.0
				66	8.5

TABLE II

A TABLE FOR CALCULATING T-SCORES
(N > 24 < 51)



Note: N is encircled figure.

This table should prove to be of value to college instructors and classroom teachers who desire to normalize distributions of scores of classroom tests and report scores as normalized T-scores. This table may be extended indefinitely, but it is felt that the maximum N of 50 will suffice for the great majority of classroom situations.

REFERENCES

1. Garrett, H. E. Statistics in Psychology and Education (New York: Longmans, Green and Co., 1954), Table G, p. 433.
2. Edwards, A. L. Statistical Methods in the Behavioral Sciences (New York: Rinehart and Co., 1954), Table XI, p. 511.

JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 28, Number 3, March 1960)

ESTIMATION OF TEST RELIABILITY BY ANALYSIS OF VARIANCE TECHNIQUE*

CYRIL J. HOYT and P. R. KRISHNAIAH
University of Minnesota

Introduction

IN AN EARLIER paper, Hoyt[1] showed a method of determining the best linear estimate of the error variance appropriate for expressing an internal consistency reliability coefficient for a total test score obtained by adding the scores on items scored either one or zero. In this derivation it was assumed that the pq errors of measurement in the responses of p students to each of the q items were independent. Later investigation showed that the procedure described in this earlier paper is entirely appropriate when:

1. items are scored in more than two categories[2],
2. "items" are ratings[3], test scores[4] or other measures[5].

Further theoretical development in the area of components of variance estimation has enabled the authors to show that certain assumptions regarding the independence of the errors of measurement and of independence of main effects (individuals and items) are not necessary in order that Hoyt's 1941 formulation of the reliability coefficient estimate yield appropriate ratios of expected values of the mean squares.

When item effects are assumed to be intercorrelated (see models 5 and 6) the expected values of the two variance components (A) and (C) used in estimating the reliability coefficient are precisely equal to those obtained for models 1, 2, 3, and 4 in which independence is assumed as in the 1941 derivation.

When errors of measurement are assumed to be correlated for the responses of a given individual on different items (see models 7, 8, and 9) the expected values of both variance components A (among individuals) and C (error) are different from those obtained for the first six models. However, the ratio of (A-C) to A still provides a suitable estimate of the internal consistency reliability coefficients provided the user recognizes what he has.

The definition of fixed, mixed and random models follows:

The model equation for each of the nine lines of Table I is $X_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$ where

$$i = 1, 2, \ldots, p$$
$$j = 1, 2, \ldots, q$$

where μ, α_i, β_j are the general effect, i^{th} individual effect, and j^{th} item effect. ϵ_{ij}'s are errors which are normally distributed. This equation is suitable for all three types of models with different assumptions on the α's and β's. We may suppose that the individuals and the items are drawn randomly from their corresponding populations which contain P individuals and Q items respectively.

A. Fixed Model

When P = p, and Q = q, then the above equation represents a fixed model. In this case, the only random element is the error term ϵ_{ij}. When describing the reliability coefficient for a given test form administered to a given group of individuals the fixed model is appropriate.

B. Mixed Model

When P = p and Q $\gg$ q or when P $\gg$ p and Q = q, then the equation represents a mixed model. (The symbol $\gg$ is used to designate the relation as much greater.) This model is suitable for considering the reliability obtained as an indicator of that appropriate for either measuring a given group of people on a domain of items represented by the sample of q items or, alternatively, for generalizing to a population of P subjects for any given administration of a certain test of q items.

C. Random Model

When P $\gg$ p and Q $\gg$ q then the above equation represents a random model. This model is suitable for estimating the reliability coefficient for

*All footnotes will be found at end of article.

TABLE I

EXPECTED VALUES OF MEAN SQUARES FOR NINE MODELS

Model	(A) Among Individuals	(B) Among Items	(C) Error	A - C
1. Fixed model Individual item effects fixed Errors uncorrelated	$\sigma^2 + \frac{q}{p-1}\Sigma\alpha_i^2$	$\sigma^2 + \frac{p}{q-1}\Sigma\beta_j^2$	σ^2	$\frac{q}{p-1}\Sigma\alpha_i^2$
2. Mixed model Errors uncorrelated Item effect random Individual effects fixed	$\sigma^2 + \frac{q}{p-1}\Sigma\alpha_i^2$	$\sigma^2 + p\sigma_\beta^2$	σ^2	$\frac{q}{p-1}\Sigma\alpha_i^2$
3. Mixed model Errors uncorrelated Item effects fixed Individual effects random	$\sigma^2 + q\sigma_\alpha^2$	$\sigma^2 + \frac{p}{q-1}\Sigma\beta_j^2$	σ^2	$q\sigma_\alpha^2$
4. Random model Errors uncorrelated Item effects random Individual effects random	$\sigma^2 + q\sigma_\alpha^2$	$\sigma^2 + p\sigma_\beta^2$	σ^2	$q\sigma_\alpha^2$
5. Random model Errors uncorrelated Item effects correlated Individual effects random	$\sigma^2 + q\sigma_\alpha^2$	$\sigma^2 + p(1-\rho)\sigma_\beta^2$	σ^2	$q\sigma_\alpha^2$
6. Mixed model Errors uncorrelated Item effects correlated Individual effects fixed	$\sigma^2 + \frac{q}{p-1}\Sigma\alpha_i^2$	$\sigma^2 + p(1-\rho)\sigma_\beta^2$	σ^2	$\frac{q}{p-1}\Sigma\alpha_i^2$
7. Fixed model Errors correlated on different items for same individual	$\sigma^2[1 + (q-1)\rho] + \frac{q}{p-1}\Sigma\alpha_i^2$	$\sigma^2(1-\rho) + \frac{p}{q-1}\Sigma\beta_j^2$	$\sigma^2(1-\rho)$	$\frac{q}{p-1}\Sigma\alpha_i^2 + \rho q\sigma^2$
8. Mixed model Errors correlated as in 7 Individual effects fixed Item effects random	$\sigma^2[1 + (q-1)\rho] + \frac{q}{p-1}\Sigma\alpha_i^2$	$\sigma^2(1-\rho) + p\sigma_\beta^2$	$\sigma^2(1-\rho)$	$\frac{q}{p-1}\Sigma\alpha_i^2 + \rho q\sigma^2$
9. Random model Errors correlated as in 7 Item effects random Individual effects random	$\sigma^2[1 + (q-1)\rho] + q\sigma_\alpha^2$	$\sigma^2(1-\rho) + p\sigma_\beta^2$	$\sigma^2(1-\rho)$	$q\sigma_\alpha^2 + \rho q\sigma^2$

an infinite population of subjects or a finite one of size P, on the whole domain of Q items.

Columns 2, 3, and 4 of Table I show the expected values of the mean squares components in nine different models that are each specified by the assumed characteristics listed in column one. These nine models are all useful for specifying internal consistency reliability coefficients if X_{ij} is considered the score earned by i - th individual's response to the j - th item on an examination. In these nine models the interaction term is used as error as is usually done in cases providing only one observation per cell. By making different assumptions about the parameters in the equation the nine different models each described in successive lines of Table I arise. There are a number of other models which seem useful and for which we have not yet studied the variance components and reliability estimates.

The first six entries in the last column of Table I are the expected values of the mean squares of the true scores or the numerator of the reliability coefficient appropriate for each line of the table. The denominators of the reliability coefficients in every case are the entries in column A. In the models specified in the last three lines, the entries in the last column seem to contain more than the variance of the true scores. Hence, though these could be defined as the numerators of the reliability coefficients, they do not appear to be appropriate estimates for the variance of the true scores.

Let us consider now the first four models in which all the effects represented by the terms of the equation are independent. In these cases, one notes that the usual procedure for estimating the internal consistency reliability coefficient gives adequate estimates of the ratio of the variance of the true scores to the variance of the obtained scores. The last column of Table I gives only the numerator of the reliability coefficient obtained by taking one minus the ratio of the error mean square to the among students' mean square. The denominator of the reliability coefficient is (A) of the corresponding row or the mean square for among individuals. Rows 5 and 6 show expected values for random and mixed models, respectively, in which the item effects are correlated with one another to the same extent (ρ) for each pair of items. The column indicating the expected value of the "among items" mean square gives an expression involving this common correlation ρ between items. However, as long as the errors of measurement in the several items are considered independent of each other and independent of the item and individual effects, the usual mean square ratio for estimating the coefficient of reliability gives a numerator which is an appropriate estimate of the variance of the true scores.

In models 7, 8, and 9, however, the numerator of the usual estimate of the reliability coefficient contains a term $\rho q\sigma^2$ which is not zero unless ρ is zero. In these three models, it is assumed that the errors of measurement on pairs of items are correlated with a common co-variance for each individual.

FOOTNOTES

* Work on this paper was partially supported by the Bureau of Educational Research, University of Minnesota.

1. Cyril J. Hoyt, "Test Reliability Estimated by Analysis of Variance," Psychometrika, VI (1941), pp. 153-60.

2. Cyril J. Hoyt and Clayton L. Stunkard, "Estimation of Test Reliability for Unrestricted Item Scoring Methods," Educational and Psychological Measurement, XII (1951), pp. 756-58.

3. Robert L. Ebel, "Estimation of the Reliability of Ratings," Psychometrika, XVI (1951), pp. 407-24.

4. P. Horst, "A Generalized Expression for the Reliability of Measures," Psychometrika, XIV (1949), pp. 21-31.

5. Howard W. Alexander, "The Estimation of Reliability When Several Trials are Available," Psychometrika, XII (1947), pp. 79-99.

JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 28, Number 3, March 1960)

ON TEACHING F INSTEAD OF t

NORMAN H. ANDERSON
Yale University*

IT IS WELL known that the square of a t ratio on n df is interpretable as an F ratio on 1 and n df. For the case of one or two means then, the values of t and F are equivalents, and lead to identical probability and inferential statements. Since F is the more general, applying as well to the case of more than two means, one might well suspect that the t test could be retired from use. Because the t test constitutes such a large part of current statistical pedagogy and practice, it may be useful to consider the matter in more detail. The main considerations which seem to the writer to be relevant are summarized briefly below.[1] These considerations are applicable principally to the first few statistics courses in the applied sciences and have less relevance to more theoretical courses.

In favor of t, it can be argued that it has greater intuitive appeal. To measure the difference between two means against its standard error may be simpler to comprehend and result in easier learning than does the variance partitioning technique of F. However, as soon as it is admitted that t is insufficient to meet everyday needs, so that F must be taught anyway, this argument loses its power. It may also be suggested that whatever greater immediacy of meaning may attach to t is not entirely its natural due, but rather a product of the present organization of statistics courses.

The t test may also be preferred because of its relation to some of the problems arising in estimation, such as the construction of confidence intervals. Algebraically, of course, it makes little difference whether one proceeds from t or from F. Thus, with F, the 95 percent confidence interval is obtained using $\sqrt{s_m^2 F(.95; 1, n)}$. The interpretation of the confidence interval could then be taken on faith (and even serve usefully as the basis for the significance test for the case of one or two means). However, it must be admitted that the confidence interval idea seems easier to comprehend when it is based on the standard picture of the symmetrical t distribution. This possible advantage of t must be measured against its shortcomings in problems of significance testing which are discussed below.

A further point in favor of t stems from a danger of the analysis of variance. Ordinarily, when pair-wise comparisons among several conditions are to be made, the error mean square from the overall F analysis is used throughout. When the condition variances are reasonably equal, this procedure gives increased sensitivity since the error is estimated more precisely. If there is marked heterogeneity of variance, however, such tests may be biased, and it may be preferable to test each pair of conditions using only the data from that pair (although this does not entirely escape the difficulty). Of course, the two condition F could be used here as easily as t. Nevertheless, it must be admitted that routine application of the analysis of variance does facilitate overlooking the problem.

Against these arguments must be set a number of negative aspects of t. For situations with just one or two conditions, of course, it is a matter of statistical indifference whether t or F is used. Either yields the same inferences, as noted above. Moreover, since $t^2 = F$, there need be little difference in computational efficiency. It may be noted, however, that the standard computing rules for t, in contrast to those for F, are not efficient and are thus more liable to arithmetical error.

When more than two conditions are involved, piece-meal t testing presents certain dangers.[2] The one commonly emphasized is the increased risk of thereby obtaining "significant" differences which the data do not in fact support. Curiously enough, the t test approach also increases the likelihood of the obverse error, namely, failing to detect results which the data are quite adequate to declare real. This can happen in a number of ways, which trend tests may serve to illustrate. In investigations in which one of the independent variables is numerical, the between conditions variation can often be expected on prior grounds to reside largely in the linear, or linear + quadratic, component of the curve. Pair-wise comparisons of means may fail to reveal even a fairly strong real effect which a trend analysis F will elucidate with ease.

A more fundamental objection to the t test approach concerns the original planning of the inves-

*Footnotes will be found at end of article.

tigation. The researcher who thinks in terms of t is less likely to use as incisive and informative a design as can be had for the allotted time and effort. Even when the basic plan is sound, a poor superstructure can, and not infrequently does, prejudice the analysis and interpretation, or entail considerable extra computational cost.

Fisher's commendation of the use of factorial designs for controlling extraneous variation may be cited in illustration of this last point. Even with only two main experimental conditions, it is often desirable to counterbalance a second variable, not of great interest in itself, in order to obtain added generality. Unless set up so that the two-way analysis can be and is used, however, the counterbalancing may do more harm than good.

In a similar way, it is often possible to use as a second independent variable some individual difference measure which is correlated with the response measure under study. This procedure, essentially a "randomized blocks" factorial design, may generally be expected not only to increase the sensitivity of the experiment, but also to yield information on the organismic variables.

That the F test is a desirable piece of statistical machinery hardly needs even the incidental justification given here. But if F is to be taught at all, then it evidently follows, since F is the more general, that F alone should be taught. To require the mastery of two sets of rules where one will do is an improvident use of the student's time.

For statistics courses in which time is insufficient for both tests, the choice of F would accomplish the same ends as the choice of t. If more time were available, the elimination of t would permit the development of some of the simpler ramifications of the analysis of variance, such as two-way designs, specific comparisons, and trend tests. Moreover, teaching F from the start would set up conditions for maximal positive transfer later in learning since the computing rules for F follow a uniform and elegant logic in all situations. It might also be noted in this connection that the computational procedures for correlation and regression analysis can be subsumed under the analysis of variance format as well.

Finally, it may be questioned whether learning the t test does not produce a fair amount of negative transfer. The evidence available to the writer, admittedly anecdotal in nature, suggests that students who have learned t are reluctant to abandon this skill when necessary. Having a technique which may actually, if not always advisedly, be used, they tend to regard even workaday designs as "fancy" or "not worth the trouble at this stage of investigation." Thus, if they use the design at all, they do so only with pain, and so fail to acquire the familiarity with the new technique that would come were it regarded as just another application of the F test they already knew.

It would thus appear that t is superfluous, since F may always be computed instead; inadequate, since it does not apply as generally as does F; dangerous, since its use may increase errors of inference, as well as lead to weak experiments; and a waste of time, since teaching F will provide more transfer to later learning. On the balance, it seems fair to conclude that the t test should be struck from the curriculum.

More practical experience is needed in order to integrate the analysis of variance with beginning statistics on the conceptual level. However, it is appropriate to sketch briefly the sequence of computational routines which is indicated by the argument of this note.

1. Computation of the sample variance is done using the analysis of variance table:

Source	df	SS	MS
Mean	1	T^2/N	
Error (Variability)	N-1	$\Sigma X^2 - T^2/N$	

The entry, MS_{error}, then furnishes an unbiased estimate of the population variance. The scaffolding provided by the tabular format is unnecessary, of course, but seems to serve a useful mnemonic function. Its real benefit accrues in Step 2.

2. The test of significance for a single sample is most appropriately carried out on difference scores, or on other scores which have a mean of 0 under the null hypothesis. A single division appended to the Step 1 routine yields the relevant F ratio.

3. With two samples of difference scores, the analysis of variance table is extended one step further:

Source	df	SS	MS	F
Mean	1			
Between Groups	1			
Error	N-2			

4. It is straightforward to extend Step 3 to the case of several samples and it seems advisable to do so promptly in order to avoid the development of a superstition that the technique only works for one or two. The routine remains the same when the scores are arbitrary, rather than difference scores. It is only necessary to observe that the F for Mean is no longer meaningful or important and may be omitted from the table.

5. A simple extension of the computing routine yields the calculation of r for a single sample:

Source	XX	XY	YY
Mean	T_X^2/N	$T_X T_Y/N$	T_Y^2/N
Error	$\Sigma X^2 - T_X^2/N$	$\Sigma XY - T_X T_Y/N$	$\Sigma Y^2 - T_Y^2/N$

The square of the correlation coefficient is estimated from the entries in the Error line, $r^2 = SS_{XY}^2/SS_{XX}SS_{YY}$. The test for significance may be made by consulting a table for r or, with an additional calculation, by using F. When the calculation of r is taught this way, there is less likelihood of confusedly agglomerating the between and within samples regressions when more than a single sample is involved.

It is suggested that ΣX^2 be given for most of the homework problems. This will greatly decrease the time spent on each problem and, by thus allowing a greater variety of problems to be assigned, increase the learning efficiency. Whether or not further applications of the analysis of variance are given in the first course will depend on the available time, and on whether the course is oriented toward experimental or observational research.

FOOTNOTES

* Now at the University of California, Los Angeles.

1. For the purposes of this note, the terms "F test" and "analysis of variance" will be taken as synonymous although this is not strictly correct. It will be necessary to allude to a number of uses of the analysis of variance, but these uses are well known and will not be documented.

2. A specific comparison based on some reasonably firm rational or empirical ground may always be tested, of course. However, unless it is based on a single degree of freedom, it will require an F ratio in its own right. It should also be specifically noted that multiple comparison procedures will be superior to F (and t) in certain situations.

JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 28, Number 3, March 1960)

COMPUTATION OF VARIANCE ACCOUNTED FOR IN MULTIPLE CORRELATION

CLINTON I. CHASE
Idaho State College
Pocatello, Idaho

WHEN MULTIPLE correlation techniques are used in the analysis of data, the investigator usually wishes not only to discover a combination of variables which will yield maximum prediction of a criterion, but also he wishes to determine the amount of variance in the criterion which is associated with each of the independent variables. For this purpose Engelhart (1) has submitted the formula:

$$R^2 = \beta_{O1}^2 + \beta_{O2}^2 + \ldots \beta_{On}^2 + 2\beta_{O1}\beta_{O2}r_{12} + 2\beta_{O1}\beta_{O3}r_{13} + \ldots 2\beta_{O,n-1}\beta_{On}r_{n-1,n} \quad [1]$$

Following this procedure, the ith beta squared will produce the variance directly associated with the ith variable. But since the ith variable is usually correlated with other independent variables, it also enters into variance which is shared by several of these other independent variables. The total amount of this shared variance is given by the summation of all of the terms of the form $2\beta_{Oi}\beta_{Oj}r_{ij}$ which appear in formula [1]. Then, according to Engelhart's procedure, the shared variance is divided among each of the independent variables in the same proportions as the direct variance. The total amount of variance associated with the ith variable, then, would be the ith beta squared, plus a proportional amount of the total shared variance.

Since Engelhart's procedure for partitioning the shared variance is admittedly arbitrary, the purpose of this paper is to present an alternate procedure, which, it is believed, is somewhat less arbitrary in its method of apportioning shared variance, since here the basis for partitioning the shared variance is found in mathematical manipulations.

As a necessary part of this presentation a second technique (2:409-10) for computing the variance in the criterion associated with any given independent variable must be noted. Such a technique is provided in the formula:

$$R^2 = \beta_{O1}r_{O1} + \beta_{O2}r_{O2} + \beta_{O3}r_{O3} + \ldots \beta_{On}r_{On} \quad [2]$$

Using this procedure, one may find the total amount of variance associated with the ith variable in the product $\beta_{Oi}r_{Oi}$. Although formula [2] is simpler to use than formula [1], especially when the number of independent variables is more than two or three, it has the disadvantage of not yielding a break-down between variance which is directly associated with the ith variable, and that portion of the variance shared among the several independent variables which is to be associated with the ith variable.

It therefore seems that some technique is desirable which would retain simplicity and economy of effort, while yielding a maximum amount of information. Such a system is found in the mathematical analysis of formulas [1] and [2] as shown below.

In a set of normal equations for the least-square estimate of the betas for the multiple linear-regression equation, the r's are found as follows:

$$r_{O1} = \beta_{O1} + r_{12}\beta_{O2} + r_{13}\beta_{O3} + r_{14}\beta_{O4} \quad [3]$$
$$r_{O2} = \beta_{O2} + r_{12}\beta_{O1} + r_{23}\beta_{O3} + r_{24}\beta_{O4} \quad [4]$$
$$r_{O3} = \beta_{O3} + r_{13}\beta_{O1} + r_{23}\beta_{O2} + r_{34}\beta_{O4} \quad [5]$$
$$r_{O4} = \beta_{O4} + r_{14}\beta_{O1} + r_{24}\beta_{O2} + r_{34}\beta_{O3} \quad [6]$$

Therefore, the right hand portion of [3] can be substituted for r_{O1} in formula [2], the right hand side of [4] can be substituted for r_{O2} in [2], and so on. The result would be as follows:

$$R^2 = \beta_{O1}(\beta_{O1} + r_{12}\beta_{O2} + r_{13}\beta_{O3} + r_{14}\beta_{O4}) + \beta_{O2}(\beta_{O2} + r_{12}\beta_{O1} + r_{23}\beta_{O3} + r_{24}\beta_{O4}) + \beta_{O3}(\beta_{O3} + r_{13}\beta_{O1} + r_{23}\beta_{O2} + r_{34}\beta_{O4}) + \beta_{O4}(\beta_{O4} + r_{14}\beta_{O1} + r_{24}\beta_{O2} + r_{34}\beta_{O3}) \quad [7]$$

Multiplying to eliminate the parentheses one may find the following results:

*The author wishes to acknowledge the advice received from R. F. Jarrett of the Department of Psychology, University of California, Berkeley, California.

$$R^2 = \beta_{01}^2 + r_{12}\beta_{02}\beta_{01} + r_{13}\beta_{03}\beta_{01} + r_{14}\beta_{04}\beta_{01} + \beta_{02}^2 + r_{12}\beta_{01}\beta_{02} + r_{23}\beta_{03}\beta_{02} + r_{24}\beta_{04}\beta_{02} + \beta_{03}^2 + r_{13}\beta_{01}\beta_{03} + r_{23}\beta_{02}\beta_{03} + r_{34}\beta_{04}\beta_{03} + \beta_{04}^2 + r_{14}\beta_{01}\beta_{04} + r_{24}\beta_{02}\beta_{04} + r_{34}\beta_{03}\beta_{04} \quad [8]$$

It can now be seen that summating like terms in [8] will produce formula [1] as it would be written in a four variable problem:

$$R^2 = \beta_{01}^2 + \beta_{02}^2 + \beta_{03}^2 + \beta_{04}^2 + 2\beta_{01}\beta_{02}r_{12} + 2\beta_{01}\beta_{03}r_{13} + 2\beta_{01}\beta_{04}r_{14} + 2\beta_{02}\beta_{03}r_{23} + 2\beta_{02}\beta_{04}r_{24} + 2\beta_{03}\beta_{04}r_{3} \quad [9]$$

Therefore, from formula [3] and [7] it is apparent that the combined direct and shared variance associated with variable one, as given by the $\beta_{0i}r_{oi}$ product in formula [2], is actually equal to the beta squared, plus one-half of the summation of all the covariance terms in formula [9], of the form $2\beta_{01}\beta_{oj}r_{ij}$, which include the beta of variable one. Likewise, it can be generalized that the total direct and shared variance in the criterion associated with the ith independent variable is given by the square of the beta for the ith variable, plus half of all the covariance terms in formula [1] which include the beta for the ith variable.

It appears, then, from a mathematical viewpoint that the variance which the ith variable shares with other variables will not necessarily be distributed in the same proportion as is the variance directly associated with that variable, as Engelhart proposed. Noting the relationship between formula [1] and formula [2], one may now compute both the variance which is directly associated with the ith variable, and that portion of variance shared among the independent variables which is to be associated with the ith variable, by using the simpler procedure as given by formula [2], with the addition of a further, simple computation. First, the $\beta_{oi}r_{oi}$ product is found, yielding the total direct and shared variance associated with the ith variable. Then β_{oi} is squared, producing the variance directly associated with the ith variable. This figure is now subtracted from the $\beta_{oi}r_{oi}$ product, the difference being that portion of the variance, shared by the independent variables, which is associated with the ith variable.

It is believed that this method provides two advantages over the procedures represented by either formula [1] or [2] alone: first, it yields maximum information concerning a given variable's participation in the variance accounted for in the criterion; and second, it materially reduces the number of computations necessary to secure that amount of information.

REFERENCES

1. Engelhart, Max D. "The Technique of Path Coefficients," Psychometrika, I (1936), 287-93.
2. Guilford, J. P. Fundamentals of Statistics, 3rd Edition (New York: McGraw-Hill Book Co., 1956).

JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 28, Number 3, March 1960)

ON A TEST OF WHETHER TWO SETS OF OBSERVATIONS BELONG TO A CERTAIN CLASS OF DISTRIBUTIONS WHEN THE RANDOM VARIABLES ARE NOT INDEPENDENT

H. P. KUANG*
University of Minnesota

THERE EXIST several techniques which have been developed in recent years to test the hypothesis that two samples have been drawn from the same population. These developments were based on signs, ranks, or the randomization principle. For instance, to test the hypothesis that the two samples are from the same population, Mood considered a test based on signs while Dixon presented a statistic based on ranks. Utilizing the randomization principle, Pitman proposed the difference of sample means as a criterion, large values being significant. Wald and Wolfowitz devised the total number of runs in a sequence for testing the same hypothesis that two samples (not necessarily of equal size) have been drawn from a common population. The tests developed are essentially limited to the continuous and independent variables. We shall, however, treat the problem without these restrictions. It is the purpose of this paper to present a probabilistic inequality which may be used to determine whether two variables are contained in a certain set of distributions without the assumptions of continuity and independence.

Let us consider the two random variables X and Y which are not independent. Suppose that each of the variables is divided into a finite number of parts A_1-----A_r, and let the corresponding discrete marginal distribtuions of X and Y be $F_1 = \{p_i\}$ and $F_2 = \{q_i\}$ $(i = 1 \text{ -- } r)$ respectively, so that $p_i = P(X = A_i)$, $q_i = P(Y = A_i)$ and

$$\sum_{i=1}^{r} p_i = 1, \quad \sum_{i=1}^{r} q_i = 1 \quad .$$

Let, further, the corresponding observations X and Y be $\{n_i\}$ and $\{m_i\}$ $(i = 1 \text{ --- } r)$ respectively, so that the sample values n_i and m_i belong to the set A_i and

$$\sum_{i=1}^{r} n_i = n, \quad \sum_{i=1}^{r} m_i = m.$$

Any set of A_i carries the mass $\frac{n_i}{n}$ in the distribution of X and the mass $\frac{m_i}{m}$ in the distribution of Y. The object here is to decide on the basis of two sets of observations $\{n_i\}$ and $\{m_i\}$ $(i = 1 \text{ --- } r)$ whether the two distributions F_1 and F_2 belong to the same class of distributions.

Using the Minkowski inequality, the author has proved in his unpublished dissertation (University of Minnesota) that if the two distributions F_1 and F_2 belong to a certain set of distributions, the following probabilistic inequality must hold:

$$P\left\{\Phi(F_1, F_2) \geqslant 1 - \frac{\epsilon^2}{2}\right\} > 1 - \frac{r^2 + r - 1}{\epsilon^4}\left(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{m}}\right)^4$$

where $\Phi(F_1 F_2) = \sum_{i=1}^{r} \sqrt{\frac{n_i}{n}\,\frac{m_i}{m}}$ and ϵ is any positive numerical value. It should be noted that the derivation of the inequality was based on the conditions that $n > r$, $m > r$, and that when $n_i \leqslant 1$ or when $m_i \leqslant 1$, $i \neq 1$. By use of the above result we can then evaluate an upper bound of the probability in terms of a pre-assigned risk, α. It is easily seen that if F_1 and F_2 have the same distribution, we get

$$\Phi(F_1 F_2) \geqslant 1 - \tfrac{1}{2}\sqrt{\frac{r^2 + r - 1}{\alpha}}\left(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{m}}\right)^2 \quad .$$

As a numerical illustration, consider the two

*The author is currently Professor of Mathematics at The North Carolina Agricultural and Technical College in Greensboro, North Carolina.

distributions of grades which were received by students enrolled in courses in calculus and differential equations at a university. The number of students assigned to each of five grades A, B, C, D, and F for these two courses were as follows:

	A	B	C	D	F	Total
Calculus	12	31	52	57	48	200
Differential Equations	2	20	44	50	44	160

The calculus was a prerequisite for the course in differential equations. These two courses were composed of not only the same individual students but also some twins, brothers and sisters. Very few of these students were, however, transfers. It frequently happened in mathematics courses that students with a good record in calculus usually made a good grade in differential equations. On the other hand, students with a poor grade in calculus were ordinarily unable to do satisfactory work in differential equations. For these reasons, the two distributions of grades are not independent. Accordingly, it is difficult to see how the ordinary test based on the assumption of independence could possibly be used. We now wish to know whether the two distributions of grades have come from the same population without the customary restriction of independence. The table at the bottom of this page gives the numerical calculations of $\phi(F_1 F_2)$.

From the last row of the table, we obtain

$$\phi(F_1 F_2) = \sum_{i=1}^{5} \sqrt{\frac{n_i}{n}\,\frac{m_i}{m}} = .989 \ .$$

If we use a risk, α, of .05, say, we get

$$\epsilon^2 = \sqrt{\frac{25+5-1}{.05}}\left(\frac{1}{\sqrt{200}} \ \frac{1}{\sqrt{160}}\right)^2 = .5395 \ .$$

Thus $\phi(F_1 F_2) \geq 1 - \frac{\epsilon^2}{2} = .730$.

Therefore, the two distributions of grades in question belong to the same set of distributions.

Finally, it may be added that if the random variables X and Y are assumed to be independent, we then have

$$P\left\{\phi(F_1, F_2) \geq 1 - \frac{\epsilon^2}{2}\right\} \geq 1 - \frac{16(r^2 + r - 1)}{\epsilon^4}\left(\frac{1}{n^2} + \frac{1}{m^2}\right) + \frac{256(r^2 + r - 1)^2}{\epsilon^8\, n^2\, m^2}$$

when $F_1 = F_2$.

It can be readily shown that the previous result (without the assumption of independence) is a better bound of probability when the following inequality holds:

$$15(n^2 + m^2) - 4\sqrt{nm}\,(\sqrt{n} + \sqrt{m})^2 > \frac{256(r^2+r-1)}{\epsilon^4}$$

	A	B	C	D	F
Calculus $\frac{n_i}{n}$	.060	.155	.260	.285	.240
Differential Equation $\frac{m_i}{m}$	.013	.125	.275	.312	.275
$\sqrt{\frac{n_i}{n}\,\frac{m_i}{m}}$	.028	.139	.267	.298	.257

REFERENCES

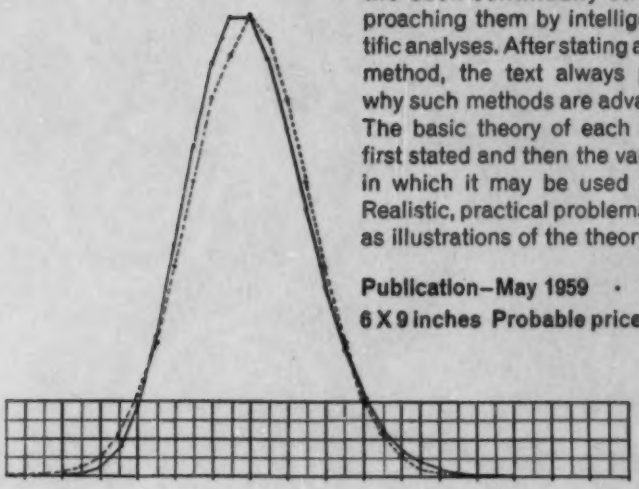
1. Dixon, W. J. "A Criterion for Testing the Hypothesis that Two Samples are from the Same Population," Annals of Mathematical Statistics, XI (1940), pp. 199-204.
2. Matusita, K. "Decision Rules Based on the Distance for Problems of Fit, Two Samples and Estimation," Annals of Mathematical Statistics, XXVI (1955), pp. 631-40.
3. Mood, A. M. Introduction to the Theory of Statistics (New York: McGraw-Hill, 1950), pp. 390-95.
4. Pitman, J. G. "Significance Tests Which May be Applied to Samples from Any Population," Supplement Royal Statistical Society Journal, IV (1937), pp. 119-30.
5. Wald, A. and Wolfowitz, J. "On a Test Whether Two Samples are from the Same Population," Annals of Mathematical Statistics, XI (1940), pp. 147-62.

MODERN STATISTICAL METHODS: DESCRIPTIVE AND INDUCTIVE

by **PALMER O. JOHNSON**
University of Minnesota
and **ROBERT W. B. JACKSON**
University of Toronto

This book is written for students who wish a functional knowledge of statistics to help them adequately solve problems needing statistical analysis.

The text not only relates how the various statistical techniques work, but also why the techniques are used and how their properties are derived. When interpreting statistical methods, the book continually stresses approaching them by intelligent, scientific analyses. After stating a statistical method, the text always points out why such methods are advantageous. The basic theory of each method is first stated and then the various ways in which it may be used are given. Realistic, practical problems are given as illustrations of the theory in use.



Publication–May 1959 · 525 pages
6 X 9 inches Probable price $7.50 list.

College Department
RAND McNALLY & COMPANY
P. O. Box 7600, Chicago 80, Illinois